fhi360
THE SCIENCE OF IMPROVING LIVES

EDUCATION POLICY AND DATA CENTER
Making sense of data to improve education

Methodology for Evaluating Data Quality

By Laurie Cameron

WP-07-02

In 2011, FHI acquired the programs, assets, and expertise of AED.

Methodology for Evaluating Data Quality

Working Paper WP-07-02

By Laurie Cameron
July 2005

Working Papers disseminated by the EPDC reflect ongoing research and have received limited review. Views or opinions expressed herein do not necessarily reflect the policy or views of the EPDC, of FHI 360, or any of the EPDC sponsors.

Recommended Citation
Cameron, Laurie, 2005. Methodology for Evaluating Data Quality. Working Paper WP-07-02. Education Policy and Data Center, Washington, DC (FHI 360)

EPDC
The Education Policy and Data Center (EPDC) was launched in 2004 with funding from the U.S. Agency for International Development, an associate award under the Education Quality Improvement Program 2: Policy, Systems, Management (EQUIP 2), to increase the breadth and depth of information easily accessible to researchers and practitioners in international education.

Table of Contents

Tables

Figures

Diagrams

BACKGROUND

Professionals in the international education community need to be able to rely on the statistics they use as a basis for programming, policy making, monitoring, and evaluation. The statistics available to them are of extremely varying quality – some are excellent, within a few percentage points of the actual value; some are fair; some are poor; and a portion is bogus. This situation, with statistics of low quality, or worse yet, of unknown quality, is detrimental in that it discourages the use of evidence for decision-making, or leads to decisions based on erroneous information.

Education statistics seldom come with an evaluation of their quality. Individual professionals thus end up relying on their own or colleagues' expert judgment, triangulation, or personal rules of thumb to determine the quality of numbers at their disposal. Each professional uses different rules and methods, leading to diverging evaluations of the statistics. In addition, there is no accumulation of knowledge on good evaluation, and no development towards a standard practice. The situation is an impediment to agreement on the facts, clear communication, on-time data delivery, and accurate decision-making. It thus compromises the quality of international education programs and education systems.

This paper provides a methodology for assessing the quality of education data by identifying their underlying factors. Such factors have been identified in the literature in many disciplines concerned with census and sample surveys. These same disciplines also provide methodologies to address the issues, which range from quality control to sophisticated statistical adjustments. In principle, the statistical or administrative institutions, which have full access to the micro data and are in possession of good auxiliary information, are best positioned to address such issues. In reality, the extent to which such methods are applied depends on the resources available to the institution collecting those data, which in turn depends on capacity, competing priorities and political will.

The goal of the assessment is to push the envelope concerning the quality of data in the education sector in developing countries by closing the gap between the knowledge about the underlying factors affecting their reliability and efforts towards their improvement. The first step is to identify the source(s) of data underlying a given statistic. The second step is to identify factors that affect the quality of the data. The third is to develop methods for assessing these factors. Using the tools identified in the third step, two parallel purposes can be achieved. The first is to assess the quality of existing data systematically. The second is to advocate for their ongoing improvement.

To achieve the first purpose, the AED's Education Policy and Data Center, hereinafter referred to as the data center, will adopt a methodology to assign a "data quality index" to selected education indicators commonly used by policy makers and other managers in the education sector. Users of the index should bear in mind its intent and its limitations. First, it is an arms length assessment. The data center will use its specialized resources to obtain information to the extent possible, together with a systematic approach, to make a judgment about data quality. Such resources include the establishment of a network of country and international experts to provide the necessary metadata and contextual information for the assessment. However, for

much of the data, the data center does not have full access to the micro data. The assessment therefore is a blunt instrument providing an overall impression of reliability, not a finely tuned statistic. Second, the methodology is experimental. It is only by putting it to the test in a variety of countries and from a variety of sources that its effectiveness as an assessment tool can be determined.

The second purpose—advocating for improvements in data quality—is achieved in part by the very effort to assess it. Educationists have long recognized that “you get what you measure”. The use of a systematic methodology by the center to measure data quality in the education sector establishes a precedence for it.

The paper consists of four sections. In the first section, the indicators for which the data reliability indexes are assigned are identified and defined. This is followed by a brief discussion of the most common data sources for these indicators. The second section comprises an inter-disciplinary review of the literature—how data reliability is defined and evaluated. The third section outlines the methodology to be applied. This methodology comprises a set of evaluation criteria drawn from the literature to identify data issues and quantify them to the extent possible. In the fourth section, the methodology is tested on a set of indicators derived from the most common data sources as identified in section I.

Methodology for Evaluating Data Quality

1. EDUCATIONAL INDICATORS AND DATA SCOURCES

1.1 Indicators

The methodology developed in this paper to assess the reliability of educational data focuses on four indicators. They are as follows:

- gross or net enrollment rate,
- repetition rate,
- completion ratio, and
- student teacher ratio.

The net enrollment ratio is the number of children of the official age group enrolled in the school system for a specific year divided by the corresponding number of children in the general population of that age. The gross enrollment ratio is total enrollment in a specific level of education, regardless of age, divided by the eligible official school-age population corresponding to the same level of education in a given school year.

The repetition rate measures the extent to which students are required to repeat a grade, whether for reasons of lack of academic attainment, or due to earlier under-aged enrollment, or because of insufficient student places in the next higher grade, or level, schooling. It is calculated as the proportion of pupils enrolled at a given grade in a given school year who study in the same grade in the following year.

A number of indicators have been used to measure the completion ratio. The World Bank has defined it as enrollment in final primary grade, minus the usual number of repeaters in that grade, divided by the population of final grade age.

The pupil-teacher ratio measures the average number of students taught by each teacher in each grade in the primary and secondary sectors for a particular year. For this purpose, numbers of students and numbers of teachers are expressed in full time equivalent units. Thus it is a composite indicator drawn from the indicators that express actual enrollments by grade and number of teachers by grade.

1.2 Data and Reporting Sources

A number of organizations collect and report education data. Some reporting sources are secondary sources that rely on other organizations to provide them with primary source data while others collect and report their own primary data. Primary source data consists of the following:

- administrative data systems

- sample surveys, and
- population censuses and population estimates for non-censal years.

Reporting organizations include the following:

- government ministries,
- donor organizations, and
- private and public research organizations.

Data reliability depends on both the reliability of the primary data, and in the case of secondary reporting sources, their processing and interpretation of primary data.

Table 1.1 indicates the sources of data currently available for 20 countries on the data center's website. All reporting sources in this table report their own primary source data except the GED. The GED site contains data from the DHS and the UNESCO statistical information system, or the UIS. The UIS solicits data from national official data sources. Reporting sources are discussed briefly below.

Ministries of education routinely collect administrative data at the school level. They generally require the school to report the number of children enrolled by age, sex, and grade, and the number of children repeating a grade by sex and grade. This reporting mechanism also includes the collection of teacher data, which can take more than one format. At a minimum, schools are asked to report on the number of teachers by grade and sex. A complete enumeration of teachers with information about qualifications, age, appointment date, salary and post may be included. Such data may also come from personnel records.

The DHS, or demographic and health survey, is a nationally representative sample survey with the objectives to provide information on fertility, family planning, child survival and health of children. Within this context, the DHS provides estimates of population by age and sex as well as fertility and mortality, elements required for population projections. Data collected in the DHS relating specifically to education include information on educational attainment among household members, literacy among men age 15-54 and women age 15-49, and school attendance rates (Macro, 1996)

The DHS may include an education add-on, the EdData Survey (DES). The DES focuses on factors influencing household decisions about children's schooling including information on reasons for over-age first-time enrollment in school, never enrolling in school, and dropping out, the frequency and reasons for pupil absenteeism, household expenditures on school, etc (Macro, 2002).

Population censuses are typically conducted in developing countries at ten-year intervals and most countries now have at least two censuses. Between censuses, population projections and estimates are made either using estimates of or assumptions about fertility, mortality, and migration, or direct counts of births and deaths from vital registers. Vital registers, commonplace in developed countries, are rare in developing countries (Fosu, 2001).

Table 1.1--Reporting sources for the 20 countries on EPDC website

Country	DHS	EdData	MOE	Local govt stat office	GED	MICS	Census	Other
Guinea	x	x			x			
Ghana	x	x			x			
Senegal					x	x		
Uganda	x	x	x		x		x	
Zambia	x	x		x	x	x	x	
Tanzania	x				x			
Ethiopia	x				x			
Philippines	x				x	x		x
Indonesia					x	x		x
Jordan	x				x			x
Yemen					x			
Pakistan					x			
Honduras					x			
Nicaragua	x				x			
Bolivia	x				x	x		
Kenya	x				x	x		
Armenia	x				x	x		
Mongolia					x			
Sri Lanka					x			
Lesotho			x		x	x		

The UIS database presents data collected from some 200 countries. Questionnaires are sent to national authorities (ministries of education, ministries of finance, etc.) to be completed by national experts. Completed questionnaires are entered into the UIS database. According to UNESCO, there is a constant two-way dialogue between the Institute's team and national experts to cross-check and analyse all data that are provided by the national bodies. National statistical or educational publications are used to cross-reference figures, as well as to ensure that no change has occurred in the structure of the country's education system since the last questionnaire administered. If any inconsistencies are noted in the data presented by the national authorities, the UIS then contacts the country with questions for clarification. Finally, statistics and indicators are calculated using UN population data and finance data from the World Bank (UNESCO, 2003).

The Multiple Indicator Cluster Survey (MICS) was developed by UNICEF to provide countries with a survey methodology using probability sampling to assess progress on all goals of the World Summit for Children at the end of the decade. Through the surveys, data were collected on nutrition, health and education, as well as on birth registration. family environment, child work and knowledge of HIV/AIDS. The end-decade MICS were conducted in 66 countries, primarily by national government ministries with support from a variety of partners, to help fill many gaps in data on children in the developing world (UNICEF, 2004).

Other sources of data include special studies, data arising from project implementation, and surveys of schools or households at a regional or district level. Many of these are targeted at a subpopulation to analyze a specific issue. Such studies provide supporting evidence for the national-level indicators.

The numerator and denominator for an indicator may come from the same source or data collection activity, or may depend on separate sources. When the numerator and denominator are from different sources, data quality must be evaluated for each source. Table 1.2 indicates the sources for the numerator and the denominator of each indicator by reporting mode and data source. It encompasses the most common sources reported at the national level, but is not all-inclusive.

Table 1.2—Indicators and Data Sources

Indicator	**Reporting Source**	**Data source-Numerator**	**Data source-Denominator**
GER/NER	Official government sources[1]	School-level administrative data	Census, postcensal or intercensal estimates
	Sample survey	Person-level survey response	Person-level survey response
Repetition rate	Official government sources	School-level administrative data, current year	School-level administrative data, previous year
Completion rate	Official government sources	School-level administrative data	Postcensal or intercensal estimates derived from census data
	Sample survey	Person-level survey response	Person-level survey response
Pupil-teacher ratio	Official government sources	School-level administrative data or other administrative sources	School-level administrative data

[1] Includes other organizations who rely on official government statistics for their source.

2. LITERATURE REVIEW

Literature on data quality generally takes one of three approaches. The first approach is concerned with insuring the accuracy of data through field testing and quality control during the process of collecting information in the field, entering the data into electronic format, and processing it for dissemination. These issues relate to the production process.

A second approach is concerned with assessing the accuracy of data once it has been collected and processed. These types of evaluations are conducted not only for the intrinsic value of such information, but also to determine whether adjustments need to be made, and to improve the quality of data in the next round of data collection. A large body of literature in this area is in the field of censuses. Because of the prohibitive cost of producing a census, they are usually only conducted every ten year, hence lessons learned from one census play a pivotal role in designing the next census. Additionally, the census provides the basis for population estimates in non-census years. Therefore the reliability of the census is critical.

Whereas the first two approaches are concerned with the accuracy of data, the third approach encompasses a broader array of issues relating to the overall quality of data such as relevance, accessibility, integrity, and interpretability (defined below). This approach was embraced by the IMF in its initiative to establish standards in the dissemination to the public of its members' economic and financial data. More recently, the Data Quality Act passed by the U.S. Congress in 2001 sought to ensure the quality of information that federal agencies disseminate and use where quality encompassed such dimensions as objectivity, utility and integrity.

The focus of this paper is to develop a methodology to assess the reliability of education data with a primary focus on accuracy. But accuracy depends on or is influenced by the broader dimensions of quality. It will become evident as the methodology is developed that the ability to assess the accuracy of data depends heavily on factors like accessibility, coherence and interpretability and these factors are built into the assessment.

It is useful to start with a discussion of literature on data quality in the broadest terms. This sets the stage for the development of the methodology to assess the reliability of the data. Methodologies provided in the literature to assess the accuracy of the data are then presented in broad strokes. In section III specific methodologies are selected for the task at hand.

2.1 Defining Data Quality

The IMF began work on data dissemination standards in 1995 to guide members in the dissemination to the public of their economic and financial data. The purposes of the initiative, called the general data dissemination standards (GDDS) were to 1) encourage member countries to improve data quality; 2) provide a framework for evaluating needs for data improvement and setting priorities in this respect; and 3) guide member countries in the dissemination to the public of comprehensive, timely, accessible, and reliable economic, financial, and socio-demographic statistics.

The Fund's website[2] includes references to a number of papers relating to data quality. A paper supplied by Statistics Canada (Brackstone, 1999) summarizes the general approach to data quality by looking at six dimensions about which a statistical agency needs to be concerned:

- **Relevance** reflects the degree to which statistical data meets the real needs of the client. Is the information covering the right topics and utilizing the appropriate concepts for measurement within these topics?
- **Accuracy** is the degree to which statistical data measures what is was intended to measure. Accuracy is usually concerned with concepts of standard error, bias, and confidence intervals. It may also be described in terms of major components of potential error that may cause inaccuracy such as coverage, non-response, measurement, etc.
- **Timeliness** refers to the delay between the reference point and the date the data becomes available. Accurate information on relevant topics won't be useful to decision-makers if it arrives only after they have to make their decisions.
- **Accessibility** refers to the ease with which data can be obtained. This includes the ease with which the existence of information can be ascertained as well as form and availability of the actual data.
- **Interpretability** reflects the availability of supplementary information and metadata necessary to interpret and utilize data appropriately. To make use of statistical information, the users needs to know what they have and understand the properties of the information. Data dissemination should be accompanied by descriptions of the underlying concepts, processing and estimation used in producing the information, and its own assessment of the accuracy of the information.
- **Coherence** reflects the degree to which data can be brought together with other statistical information within a broad analytical framework or over time. Users are often faced with utilizing different sets of statistical information derived from different sources at different times. Appropriate use is facilitated if information can be validly compared with other related datasets. This facility is achieved through the use of common concepts and methodologies across time and space.

More recently, the U.S. Congress passed the Data Quality Act (DQA) in 2001. Congress enacted the DQA primarily in response to increased use of the internet, which gives agencies the ability to communicate information easily and quickly to a large audience. Under the DQA, federal agencies must ensure that the information it disseminates meets certain quality standards. Congress' intent was to prevent the harm that can occur when government websites, which are easily and often accessed by the public, disseminate inaccurate information.

Quality is defined by the OMB (2002) as an encompassing term comprising objectivity, utility, and integrity. Objectivity is a measure of whether the information is accurate, reliable, and unbiased. The federal guidelines stress the use of quality control of its production to achieve objectivity. Factors affecting the utility of data, or usefulness from the public's perspective,

[2] International Monetary Fund. http://dsbb.imf.org/Applications/web/dqrs/dqrsapproaches/

include the degree to which it is transparent, has been subjected to rigorous robustness checks, and is documented in terms of sources, quantitative methods and assumptions used. Integrity refers to the security of information to prevent it from being compromised through corruption or falsification.

2.2 Evaluating Data Accuracy

- Literature concerned with assessing the accuracy of data begins by identifying components of potential error that may cause inaccuracy. It then discusses methodologies for assessing the extent of such errors.
-

2.21 SOURCES OF ERRORS

Recall that sources of data used for educational statistics come from sample surveys, administrative data systems, and censuses. A common form of administrative data system in education is an education management information system, or EMIS. Note that most EMIS activities collect data for the entire population of schools, or a full school census.

In general, sources of errors in full censuses include the following:

- Coverage errors occur when the survey unit is missed, incorrectly enumerated, or counted more than once.
- Unit non-response errors occur when responses cannot be obtained from certain survey units.
- Item non-response errors occur when the respondent fails to respond to some of the items in the survey instrument.
- Measurement or response errors occur when the respondent misunderstands a survey question and/or records an incorrect response.
- Processing errors occur during coding, data capture, and editing.

Coverage errors affect the accuracy of census counts—populations, families, households, and dwellings. Errors occur when persons or dwellings are missed, incorrectly enumerated, or double counted. On average under-coverage is more likely to occur than over-coverage and, as a result, counts are likely to be understated. As well as affecting total population counts, undercounting can bias other census statistics because characteristics of the missed survey units are different from those that are counted (Statistics New Zealand, 2001).

Unit non-response errors occur when responses cannot be obtained from the survey or census unit for whatever reason. Like coverage, non-response can affect total counts and bias other statistics, depending on the characteristics of non-responders. Schools located in remote areas with poor communication infrastructure are more likely to fail to report administrative data. These are also schools that are likely to have lower enrollment rates and higher pupil teacher ratios.

Item non-response refers to the fact that respondents do not always respond to every question on a survey or instrument. This is likely to be true if the questions are of a sensitive nature, if the request for information is not well understood, or if the instrument is long and overly detailed.

Measurement or response errors can occur because the respondent misunderstands the question or can only approximate an answer, or simply transfers it incorrectly, for example, from school records onto the form. However, in the case of routine administrative reporting, there can also be incentives to over- or under-represent certain data, for example, if financial and other resources allocated to the school are linked to enrollment levels or performance.

Processing errors occur during coding, when “write-in” responses are transformed into numerical codes, during data capture, when the responses are transferred from paper to electronic format, and during editing, when a “valid”, but not necessarily correct, response is inserted into the record to replace a missing or “invalid” response. (Parsons, 1999).

Finally, sampling errors arise from the fact that the responses to the survey instrument, when weighted up to represent the whole population, inevitably differ somewhat from the responses that would have been obtained if these questions had been asked of the whole population (Macro, 2001).

With the exception of coverage errors, these issues also apply to sample surveys. Issues unique to sample surveys include problems when the incomplete sample coverage, and sample design and use. Incomplete sample coverage occurs when the list used to select survey samples excludes part of the population. The sample design must insure that units are selected with a known probability. Where the set of characteristics being measured varies by subpopulations, the sample design is improved by stratifying the population into subpopulations and drawing samples form each subpopulation. Subsequent analyses representative of the entire population must be weighted by the inverse of the probability of selection.

2.22 EVALUATION METHODS

Methods for evaluating data quality can be on-going during the entire process from field interviews to data processing, or can be used to evaluate a data set upon dissemination. Many of the techniques used for data quality, in addition to evaluating the data, provide methods for making adjustments. For example, a coverage rate for a census that is below 100% is not necessarily an indicator of poor data quality if appropriate adjustments have been made to it. In evaluating data quality, therefore, it is necessary to explore the sources and extent of error and what has been done to address them.

Evaluation methods described in the literature include the following:

- Field testing
- Quality assurance procedures
- Post enumeration surveys

- Non-response analysis
- Demographic methods
- Triangulation
- Estimation of variance and standard errors

The survey instrument or questionnaire is **field-tested** prior to the launch of a survey or census to reduce response and measurement errors that arise through poorly worded or ambiguous queries (Australia Bureau of Statistics, 2001). Pretesting procedures may involve cognitive interviews, focus groups, respondent debriefing, interviewer debriefing, split sample tests, and analysis of item non-response rates and response distributions (US Bureau Of Census, 1993).

Quality assurance procedures continuously evaluate data during data capture and processing to minimize processing errors (Statistics New Zealand, 2001). Such procedures include double keying, range checks, benchmarking, and monitoring internal consistency—e.g., do the individual items sum to the total, does date of birth precede data of death. In addition, editing and certain kinds of imputation are generally built into the data processing system.

The **post-enumeration survey (PES)** is a method for evaluating the results of a census. As censuses become more complicated, and as the results of censuses are used for more and more policy and planning purposes, it is important to examine the quality and limitations of census data and to understand the types and extent of inaccuracies that occur. Basically, a PES is an independent sample survey that replicates a census. The survey results are compared with census results, permitting estimates to be made of coverage and content errors. The PES allows census organizations to uncover deficiencies in the methodology of the census and make adjustments for future censuses. PES results can also be used to adjust census results, although this is as likely to be a political decision as a technical one. In many developing countries, alternative sources of population data are not available, so the PES is the major tool for evaluating the census (Whitford, 2001).

Non-response analysis is used to measure the level of non-response for both the responding unit and for individual items on the survey instrument for which responses were not obtained (U.S. Bureau of Census, 2000). Non-response is easily estimated. Reasons for non-response are more difficult to ascertain by the very fact that the units are non-responding. Some studies include efforts to identify characteristics of non-respondents though follow-up research or secondary sources to determine whether subsets of the population were systematically under-represented. Unit non-response may be adjusted for using a non-response adjustment factor in the calculation of statistics. Item non-response of critical data is often imputed using methods ranging from substitution of respondent's record from a previous cycle of the survey to hot decking, where other respondents with similar characteristics are used as "donors" of information for the missing response.

Demographic analyses rely on the fact that there is a deterministic relation between population size and structure, and fertility, mortality, and migration. It is therefore possible to project or simulate population levels, age distributions, cohort survival rates, etc., across time and compare them to actual data from periodic sample surveys and censuses (Fosu, 2001). Such

methods are also used to evaluate the quality of postcensal and intercensal estimates (Vladislaw, 1999).

Demographic methods include but are not limited to the following (Fosu, 2001, Vladislaw, 1999):

- The graphical analysis of age sex distributions (age-sex pyramid) has become the standard method of evaluating all population censuses
- The cohort component method measures deviations of a new census from population projections from a previous census.
- The cohort survival method compares the size of birth cohorts in successive censuses. In a population closed to migration, the variations in a birth cohort between two successive censuses are attributed to mortality.
- Age specific sex ratios should fall with certain ranges. Ratios outside those ranges (in the absence of extreme events) indicate likely content errors.
- Tests for linearity indicate whether intercensal estimates were obtained by linear interpolation.
- The amount of time elapsed between censuses and whether the population estimates are postcensal or intercensal provide qualitative indicators of data reliability.
- The extent to which a direct count of deaths, births, and migration are available for the years when a direct count of the full population is not available provides another qualitative indicator of the reliability of population estimates.

Note that demographic methods have varying data requirements. Moreover, the greater the availability and frequency of sources, the broader the range of demographic techniques can be used. A major weakness to many of these methods is that they do not provide enough information to separate errors of coverage from errors in content (Fosu, 1999).

Triangulation is a methodology developed by sociologists that utilizes qualitative and quantitative data collected from different sources that use different strategies to verify the accuracy of data. Since different methods entail different weaknesses and strengths, methodological triangulation consists of a complex process of playing each method off against the other so as to maximize the validity of field efforts (Udo, 2001).

Sampling error refers to the difference between the estimate derived from a sample survey and the 'true' value that would result if a census of the whole population were taken under the same conditions. “Sampling error is usually measured in terms of the standard error for a particular statistic (mean, percentage, etc.), which is the square root of the variance.” (Macro, 2001). Sampling design issues can frequently be address by increasing the size of the sample or by choosing a stratified sample. In affect, a sample is selected for each segment of the population for which responses are expected to differ.

The types of errors, the evaluation method, and the adjustment method are summarized in table 2.1. It is evident that the evaluation tools can be used for more than one type of error. In fact, many of the tools reveal data quality issues, but not necessarily their source. It is often more useful, therefore, to dichotomize sources of error into coverage and content. Even then, some

evaluation techniques will not distinguish between coverage and content. Note, finally, that in this section methodologies are discussed generally. In the following sections, specific methodologies appropriate to the type of data source being evaluated will be identified.

Table 2.1 —Type of issue, evaluation method, adjustment/correction method

Type of issue	**Quantitative method**	**Adjustment/correction method**
Coverage	Post enumeration surveys, demographic analysis, triangulation	Census coverage adjustment
Unit non-response	Unit non-response analysis, triangulation	Non-response adjustment
Item non-response	Field testing, item non-response analysis, triangulation	Imputation
Measurement or response errors	Field-testing, post enumeration surveys, demographic analyses, triangulation. variance analysis	Imputation
Processing errors	Quality assurance procedures, post enumeration surveys, demographic methods, triangulation	Editing, imputation
Sampling errors	Standard errors	None

3. METHODOLOGY FOR CONSTRUCTING DATA QUALITY INDEX

The previous section described a variety of methods and criteria that can be used for evaluating data quality. In this section, the appropriate choice for each data source is developed. Clearly the level of detail of education data, its format, and the amount of information underlying the indicators provided on the data center's website vary considerably from country to country. For some countries, such as Uganda, AED has been actively assisting their data collection and processing activities. Direct access to these data, its metadata[3], and an in-depth understanding of its context provides a rich source for mining and evaluating its quality. For other countries, data may comprise only aggregate statistics while in others, data may be very scanty. The methodology for evaluating data quality, therefore, is twofold: which dimensions of data quality should be the primary focus of the index and what information is required for that assessment.

In fact, all dimensions of data quality outlined in Section II play a role in the assessment. The indicators have been selected for their relevance in evaluating access to education and the quality and efficiency of the education sector. The methodology developed in this section for evaluating data quality focuses on accuracy. The ability to make the assessment hinges on the accessibility and interpretability of the data.

Decision trees that key off of the extent of accessibility and interpretability of the data provide an overriding structure for the evaluation. Depending on the outcome at each decision point, either an evaluation methodology is selected (upon which the indicator will be scored), or the indicator is assigned a default score. The scores, on a scale of 0-2 for each evaluation method, are used to construct an index and are outlined in the evaluation matrixes.[4] The score is based not only on the magnitude of the issue, but on whether any appropriate adjustments/corrections have been made.

Generally, each criterion is given equal weight. Given that the assessment is conducted at arms length, i.e. the data center generally does not have full access to the micro data, it is not possible to determine the magnitude of the effect of particular data issues on the data. As stated at the outset, the evaluation is intended to give on overall impression of the quality of the data using a standardized systematic review process.

Two critical issues in developing the methodology are 1) the identification of the organization that conducts the analysis and 2) the minimum data requirements for the analysis. Ideally, the organization responsible for data collection and dissemination is in a key position to evaluate data quality, make adjustments to the data and/or provide appropriate information for users to perform their own evaluations. To this end, an ongoing activity of the data center will be to encourage and assist countries to recognize problems that inevitably occur when collecting, processing, and analyzing data and to take thoughtful steps to deal with them.

[3] Metadata or "data about data" describe the content, quality, condition, and other characteristics of data.
[4] A scale of 0-2 was selected to capture the magnitude of the severity of the error—small, medium, large. Where the outcome is dichotomous, the score was either 0 or 2. This was done to give equal weight to each evaluation method.

3.1 Administrative Data

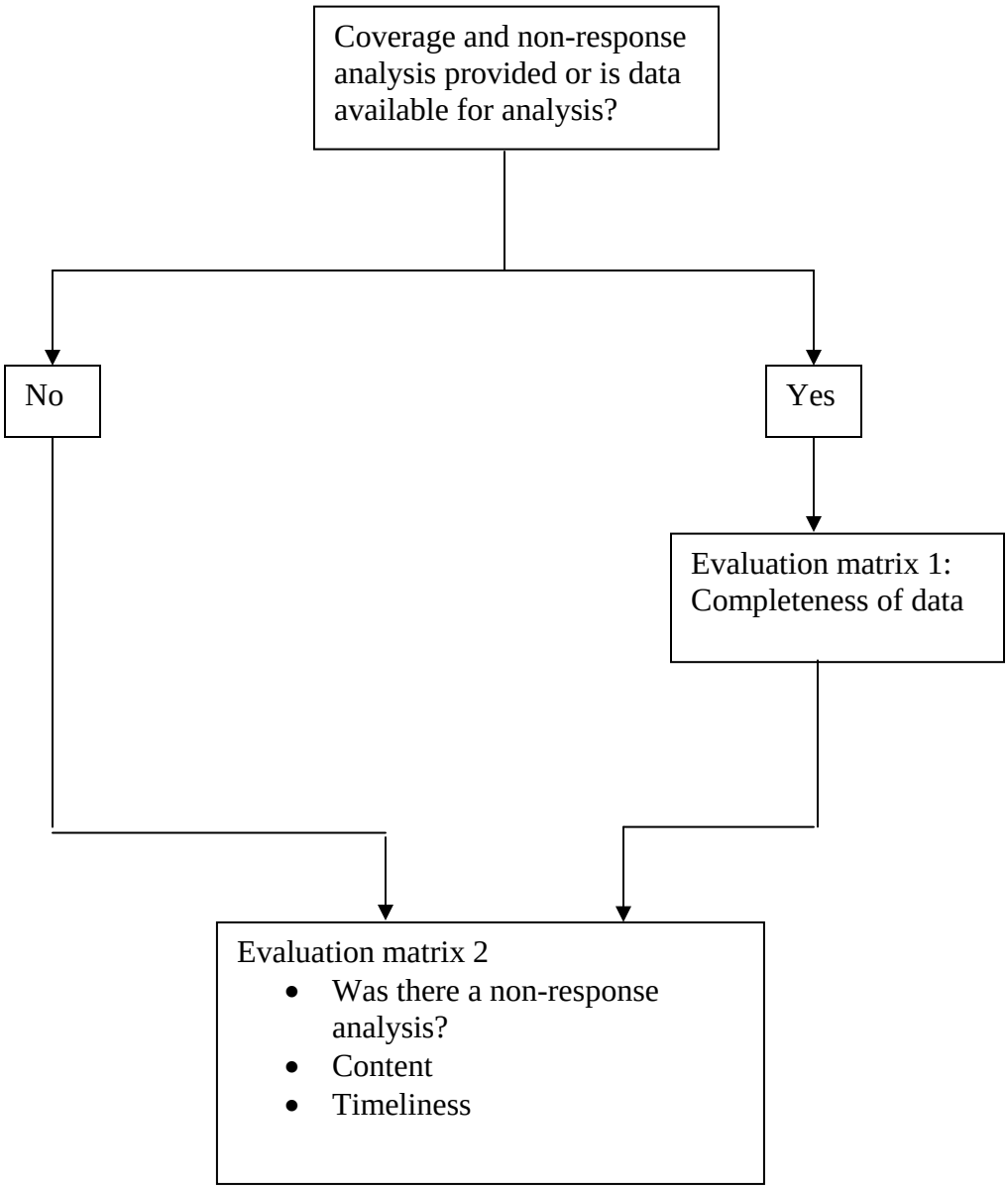
Evaluation of the quality of data as stated above focuses primarily on accuracy. The accuracy of data depends on a number of issues that are generally classified under two broad categories: completeness and content. Completeness refers to coverage and non-response, which have specific quantifiable measurements. Content refers to issues relating to data processing and measurement. Unlike methods to measure completeness, those used to identify content issues cannot necessarily be attributed to a specific issue. Moreover, methods used to evaluate content may also reveal issues of completeness.

Unfortunately, coverage and non-response, standard measures of quality reported for sample survey data and censuses, do not seem to be as common for administrative education data systems. Hence the administrative decision tree has a node for completeness: has the ministry conducted an analysis of completeness and, if not, is there sufficient data for the data center to conduct this type of analysis? The evaluation of completeness of data is followed by the evaluation of content (diagram 3.1).

An additional factor will be weighed into the assessment: timeliness. This factor is included because timely data that are used for decision-making are scrutinized and changes in trends and other anomalies are queried. People that produce data that are actively used have a greater incentive to produce quality data. The collective findings of these evaluation criteria, therefore, provide an overall impression of data quality.

The outcome of this evaluation can be used for repetition rates and pupil-teacher ratios without further analysis because the numerator and denominator derive solely from administrative data (see section I). Enrollment rates and completion ratios use population estimates from censuses and projections in the denominator. An evaluation of population estimates is thus required before the index for these indicators can be calculated.

Diagram 3.1—Decision tree for Administrative Data



The methodology that will be used to evaluate administrative data and minimum data requirements are shown in table 3.1. Details are discussed below.

Table 3.1—Evaluation method, data requirements by type of issue: Administrative

Type of issue	**Evaluation criterion**	**Minimum data requirements**
Complete-ness	Coverage: # schools surveyed/number of schools in country	Enumeration or count of all schools surveyed and estimate or count of all schools in country
	School non-response: # non-responding schools/total qualifying schools	Number of responding schools and number of schools surveyed
	Item non-response: # non-responses to question/# responding schools	Number of non-missing responses to the question and number of responding schools
Content	Enrollment trend: Percent change in enrollment per year that is explained vs. unexplained	Consecutive years of data on enrollment
	Completion trend: Percent change in enrollment in final primary grade per year that is explained vs. unexplained	Consecutive years of data on enrollment in the final primary grade
	Teacher trend: Percent change in teachers per year that is explained vs. unexplained	Consecutive years of data on teachers
	Repetition trend: Percent change in repetition per year that is explained vs. unexplained	Consecutive years of data on repetition
	Internal consistency in relationship between enrollment, repetition, promotion and dropout by grade	Two year of consecutive data on enrollment and repetition by grade
Timeliness	Number of years elapsed since data was collected	Number of years elapsed from the time data was collected until it was disseminated

3.11 EVALUATION MATRIX 1: COMPLETENESS OF DATA

Table 3.1 shows that coverage and non-response rates are measured as simple ratios requiring counts of the number of schools in the country, the number identified by the ministry and sent a survey instrument, and the number responding. Since schools that have not been enumerated are likely unknown, ministries are less likely to be able to provide coverage rates than school non-response rates. Schools that respond to the questionnaire may not respond to all

items in the questionnaire, e.g., repetition and number of teachers. Repetition and teacher non-response rates are measured from school-level data.

If a missing response is treated as a zero in the ratios, then the ratios will be under- or overstated. To address this issue, coverage and non-response affecting critical variables should be handled by imputing the missing values or applying an appropriate adjustment factor. Finally, coverage and non-response are an issue not only because undercounting affects totals, but such errors can also bias other statistics because characteristics of the missed schools or missed items are different from those that are counted. Although the bias cannot be measured, it is a fairly safe assumption that high non-response rates introduce bias in other statistics.

The evaluation matrix for completeness is shown in table 3.2a. Scores for these evaluation criteria are designated on a scale of 0-2 depending on the magnitude of the coverage or non-response as well as what adjustments have been made for them. Coverage and non-responses that have been appropriately adjusted for or imputed get full points. In the absence of adjustments, the higher the magnitude of the issue, the less points allocated to the indicator for these criteria.

Table 3.2a—Evaluation Matrix 1: Administrative Data

Evaluation matrix 1: Coverage and response analyses					Score	Total
Was coverage error measured/can it be measured?	No				0	
	Yes	Coverage error is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Coverage error is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Coverage error is less than 5%			2	2
Was school non-response measured/can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2
Was repetition non-response rate measured/ can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2
Was teacher non-response rate measured/ can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2

3.12 EVALUATION MATRIX 2: CONTENT AND TIMELINESS

Content issues are evaluated by looking at annual trends in the data that comprise the numerator coupled with qualitative information about the country and a consistency check relating to student flows through the system. Annual trends in total enrollment, repetition, the number of teachers, and enrollment in the final primary grade are examined for anomalies. Such anomalies may be indicators of a number of data issues ranging from processing errors to measurement errors to changes in methodologies across time. They may also reflect deliberate national policies or other exogenous shocks impacting on the education system. Thus, it is necessary to review the data in the national context, e.g. what policies has the government pursued to increase enrollment, improve the pupil-teacher ratio, effect repetition, etc., and have there been any exogenous effects on education reporting or trends, e.g., internal strife, natural disasters, wars. In the absence of other reasonable explanations, large variations in trends are attributed to content issues.

The cohort flow analysis is used to check for internal consistencies in the data. Using enrollment and repetition data by grade for successive years, cohort student flow is constructed to identify unlikely patterns in promotion, repetition and dropout. Specifically, the number of dropouts, which is calculated as a residual, cannot be negative.

Timeliness is measured by the number of years elapsed from the time the data was collected until it was disseminated to decision-makers.

The evaluation matrix for content and timeliness is shown in table 3.2b. As with the matrix for completeness, scores for these evaluation criteria are designated on a scale of 0-2 depending on the magnitude of the issue identified. Note that countries lacking information about the completeness of data by-pass evaluation matrix 1 and their total possible points are based only on evaluation matrix 2. They could presumably score very well on the second evaluation, but have very low response rates, so exclusion of evaluation matrix 1 ignores the importance of non-response. On the other hand, if the indicator received zero points out of a possible 8 points by including evaluation matrix 1, then too much weight would be placed on coverage and non-response. To offset the effect on the total score of the lack of a non-response analysis, therefore, administrative data for which there is no information on coverage and non-response are penalized only 2 points.

Table 3.2b—Evaluation Matrix 2: Administrative Data

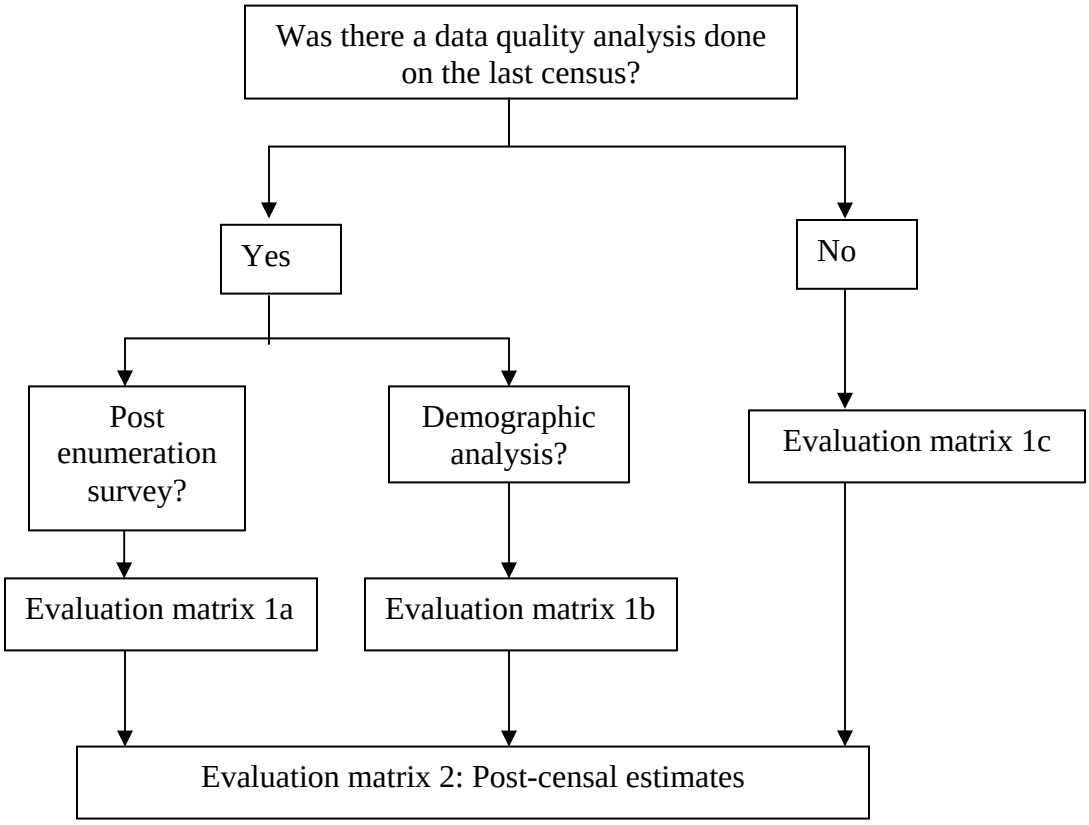
Evaluation matrix 2					Score	Total
Coverage and response analysis bypassed (if no, skip this criterion)					0	2
Content	Are there at least four years of sequential data for net enrollment?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years of sequential data for repeaters?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years of sequential data for teachers?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years data for enrollment in the final primary grade	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are enrollment and repetition by grade available for two consecutive years?	No			0	
		Yes	Are all residual dropout rates positive?	No	1	
				Yes	2	2
Timeliness	Number of years elapsed from the time the data was collected until it was disseminated			More than two years	0	
				Two years	1	
				One year	2	2

3.2 Censuses and population estimates

Population estimates are use pervasively in socio-economic analyses for policy and decision-making. Therefore accuracy is critical. At the same time, censuses are extremely costly undertakings hence they are usually only undertaken at ten-year intervals. During non-census years, projections must be made using the latest census as the base. Because the accuracy of the projection depends critically on the accuracy of the base, the evaluation of the accuracy of censuses has become standard in developed countries and is increasingly being undertaken in developing countries. Evaluations of censuses are conducted by means of a post enumeration survey and/or demographic analyses. Whereas a post enumeration survey must be conducted by the country shortly after the census, demographic analyses may be conducted by the country or the data center.

The evaluation methodology developed here examines first, the quality of the census data upon which the population estimate derives, then evaluates quality of the post-censal estimates. The evaluation of census data will utilize the results of data quality analyses conducted by the country. Where such analyses are not available, the data center will conduct demographic analyses to the extent that data are available. The decision tree for the evaluation of censuses and post-censal estimates is provided in diagram 3.2.

Diagram 3.2—Decision Tree for the Evaluation of Census and Post-censal Estimates



The methodology and minimum data requirements that will be used to evaluate census data is shown in table 3.3. Details are discussed in the following sections.

Table 3.3—Evaluation method, data requirements by type of issue: Census and post-censal estimates

Type of issue	Quality evaluation criterion	Quantitative data requirements
Census coverage and content [1]	Demographic analyses and/or post enumeration survey undertaken by country	Report by country
	Age-sex pyramid	Population distribution by single age and sex
	UN age-sex index	
Post-censal	Number of years since last census	Year of last census
	Methodology used for projection	To be provided by reporting entity
	Major shocks affecting the population	Qualitative information about the country that may have affected migration or death rates.

[1] Minimum requirements to evaluate census quality is either a report undertaken by the country's statistical body or an age and sex distribution, but not both.

3.21 EVALUATION MATRIX 1: CENSUS

Coverage and content may be evaluated by means of a post enumeration survey or demographic analyses. Coverage errors comprise under- or over-counting, with undercounting being the most common. Content issues of concern here relate to the age distribution of the population by sex. If issues are large, corrections and adjustments should be made before census estimates are used as the basis for projections.

The evaluation matrix where the country has undertaken a post enumeration survey is shown in table 3.4a (matrix 1a). In a post-enumeration survey, the extent of coverage errors—overall and by age—can be directly measured. Undercounting commonly occurs in the 0-4 and 5-9 age groups. Under-enumeration of these age groups has a significant impact on population projections, particularly school-age population projections. Therefore the evaluation matrix looks at coverage rates for the overall population and for the youngest age categories. If coverage errors are very low (less than 2%) or if adjustments have been made, full points are assigned for this criterion. If coverage errors appear to be an issue and adjustments have not been made, then a value of 0 or 1 is assigned to this criterion, depending on the level of coverage error.

Demographic analyses of census provides an alternative method for evaluating its accuracy. The evaluation matrix where the country has undertaken demographic analyses is shown in table Matrix 1b. There are a variety of demographic analyses that may be

used by a country. Such analyses may include graphical and statistical evaluations. At a minimum, they should include an evaluation of the extent of age misreporting and an evaluation of population growth compared

Table 3.4a—Evaluation Matrixes for Census Data

Evaluation Matrix 1a: Post Enumeration Survey Conducted by Country			Score	Total
Coverage	Coverage error greater than 5%	Adjustments made	2	
		Adjustments not made	0	
	Coverage error between 2 and 5%	Adjustments made	2	
		Adjustments not made	1	
	Coverage error less than 2%		2	2
Content	Coverage issues greater than 5% in age groups below 10 years	Adjustments made	2	
		Adjustments not made	0	
	Coverage issues between 2 and 5% in age groups below 10 years	Adjustments made	2	
		Adjustments not made	1	
	Coverage issues less than 2% in age groups below 10 years		2	2

Evaluation Matrix 1b: Demographic Analysis Conducted by Country					Score	Total
Age-sex analysis	Age misreporting found?	No			2	
		Yes	Adjustments made		2	
			Adjustments not made	Some irregularities	1	
				Distinct preferences for particular digits	0	2
Cohort component analysis	Population growth consistent with projections based on known fertility, mortality and migration patterns?	Difference in the two estimates greater than 5%		Adjustments made	2	
				Adjustments not made	0	
		Difference in the two estimates between 2 and 5%		Adjustments made	2	
				Adjustments not made	1	
		Difference in the two estimates less than 2%			2	2

Evaluation Matrix 1c—Census evaluation conducted in-house				Score	Total
Is age-sex distribution available?	No ⟶ go to evaluation matrix 2			0	
	Yes	Is graphical analysis of age-sex distribution consistent with expectations?	Yes	0	
			There are some irregularities	1	
			There are distinct preferences for particular digits	2	2
		UN age-sex index	Over 40	0	
			Between 20 and 40	1	
			Under 20	2	2

with known fertility, mortality, and net migration rates. Evaluation criteria are based on these reported findings. Where an issue is identified, the score assigned for the criterion will depend on the magnitude of the issue and the adjustments made. Note that since age misreporting may not necessarily be quantified (e.g. digit preferences observed from a graph), this indicator cannot a priori be assigned critical values.

Where the country has not undertaken its own evaluation of its census, the data center will undertake an evaluation using a graphical analysis and the UN age-sex index, as indicated in Matrix 1c. The age structure of the whole population at a given moment may be viewed as an aggregation of cohorts born in different years and can be represented graphically in an age-sex pyramid. Thus a population's age structure is a "map of its demographic history" (Arriaga, p. 13). Visual inspection of this and other graphical representations of the population reveal irregularities and single digit age preferences. A judgement must be made as to the extent of issues using these graphical tools.

The UN age and sex index provides a numerical estimate of the extent of age misreporting. Scores based on this index are derived from empirical analyses of the age and sex distribution in censuses from different developed and developing countries. The United Nations suggested that "the age and sex structure of a population will be (a) accurate if the joint score index is under 20, (b) inaccurate if the joint score index is between 20 and 40, and (3) highly inaccurate if the index value is over 40" (Arriaga, p. 23).

3.22 EVALUATION MATRIX2: POST-CENSAL ESTIMATES

During non-census years, estimates must be made using the latest census together with either 1) other direct counts of births, deaths and net migration, or 2) demographic techniques. One of the most reliable demographic methods of projecting population in countries without vital registers is the cohort component method which, in effect, "simulates how population changes according to its components of growth: mortality, fertility and migration."

Because of the importance of the reliability of the census as a population base, 40% of the index score rests on this evaluation. The rest of the score is based on the length of time elapsed since the last census, the methodology used to make the estimates, and the extent of strife and natural disasters affecting the population.

Due to the costly nature of censuses, it is inevitable that the time elapsed from census to the evaluation year will grow to 10 years or more. Unfortunately, the greater the time elapsed, the less reliable the estimate. Therefore this element is included as an evaluation criterion where the score declines as the number of years elapsed increases.

The second evaluation criterion is the method used to derive post-censal estimates. Methods for estimating populations during non-census years are based on vital

registers of births, deaths, and migration, on the cohort component method, or on simple mathematical growth rates. Estimates based on vital registers are the most reliable. Estimates derived from a cohort component analysis using reliable population base and intercensal estimates of fertility, mortality, and migration also provide reliable estimates. Both methods provide reliable population estimates by age and sex, which are required for education indicators. Simple growth rates should only be used for estimates of sub-populations (Arriaga, p 308). They are not reliable as the basis for education indicators at a national level.

The final criterion used to evaluate the accuracy of population projections are major upheavals affecting population. Events such as war, famine, and internal strife may have impacts on fertility, mortality and migration that are not readily measured or captured in vital registers. The greater the upheaval and the longer events last, the less reliable will be the population projections.

Table 3.4b—Evaluation Matrix for Post Census Estimates

Evaluation Matrix 2-- Post-census estimates		Score	Total
Score from evaluation matrix 1a, 1b, or 1c			4
Number of years since last census	1-4	2	
	5-8	1	
	More than 8	0	2
Are post census estimates based vital registers or on cohort component method?	Yes	2	
	No	0	2
Since the last census, have there been any upheavals causing unusual migration or death rates?	No	2	
	Events of a short term nature may have effected migration rates, but the effect is expected to be small	1	
	There have been major upheavals known to have impacted population	0	2

3.3 Sample Survey Data

Evaluation of sample survey data is the least involved of the three sources for two reasons: 1) the numerator and denominator come from the same source at a point in time and 2) there are fewer indicators provided by sample data. Issues affecting data quality are largely the same data as for derived from a sample data as from a census, with the exception of coverage. In addition, the estimates are subject to sampling error.

As sample surveys invariably provide unit and item non-response rates, no decision tree is necessary for this data type. The evaluation matrix is shown in table 3.5

Table 3.5—Evaluation Matrix for Sample Survey Data

Evaluation Matrix						Score	Total
Unit and item non-response	Was household/ person non-response measured	No				0	
		Yes	Non-response rate is greater than 10%	Was adjustment made?	No	0	
					Yes	2	
			Non-response rate is 5-10%	Was adjustment made?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
	Was item non-response rate measured?[5]	No				0	
		Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
					Yes	2	
			Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
Content	Is age sex distribution available?	No				0	
		Yes	Is age-sex pyramid consistent with expectations?		No	0	
					Yes	2	2
	Is attendance/ enrollment by grade by age available?	No				0	
		Yes	Percent of pupils in unlikely or impossible grade-age cell less than 3%		No	1	
					Yes	2	2
Sampling error	Was sampling error provided/can it be estimated?	No				0	
		Yes	95% confidence interval less than or equal to ± 2%			1	
			95% confidence interval greater than to ± 2%			2	2

3.31 NON-RESPONSE

Sample surveys invariably report household or person non-response. Unless the sample is self-weighting, weights must be applied to the sample to ensure that the sample is representative at the national level. Adjustments are made to these weights for unit non-response.

Item non-response, such as age, attendance, grade, etc are also typically reported in frequency distributions and cross-tabulations. For critical variables, missing values may be imputed. Alternatively, the persons with non-responses are excluded from the calculation of the ratio.

The sample survey data will be evaluated in terms of the level of unit and item non-response and what measures have been taken to adjust for such errors. This information can be obtained from the reported results and source documentation.

[5] Item non-response may include age, attendance/enrollment, grade, and school completion

3.32 CONTENT

Enrollment (or attendance) by age and grade provides estimates for the numerator for net enrollment (attendance) ratios. Internal consistency is evaluated with a cross-tabulation of grade by age. The population age-sex distribution derived from the same survey provides estimates for the denominator. The quality of the school-age population is evaluated with the age-sex pyramid, as with censuses.

3.33 SAMPLING ERROR

The sampling error is measured by the standard error of the indicator.

3.4 Computation of Indexes

Indexes are calculated for the numerator and denominator separately. This is done in the following steps:

- Work through the evaluation matrixes presented above to derive findings.
- Apply scoring system from findings.
- Calculate scores for each education indicator as a percent of the total possible score.
- Where numerator and denominator derive from different sources, calculate the average percentages of the two scores.
- Assign a preliminary index value based on the range in which the score falls.
- Compare value of indicator with all other sources, or triangulate across sources.

A data quality index will be assigned based on the ratio of points assigned to an indicator to the total possible (expressed as a percent) as follows:

Table 3.6—Mapping of index to evaluation scores and its interpretation

Percent of Total Possible	Index	Meaning
0-40%	1	Unreliable: unlikely or impossible values or insufficient information to rank indicator
41-60%	2	Reliability is questionable, cannot be substantiated
61-80%	3	Fairly reliable; trend analysis and cross-country comparisons require caution
81-100%	4	Highly reliable, suitable for trend analysis and cross-country comparisons.

3.41 APPLICATION OF SCORING SYSTEM

For administrative data, a number of quality assessment criteria are used. Some of these criteria apply to the overall quality of the data and so should be used for each education indicator, e.g., school response rates. Similarly, since enrollment is in either the numerator or the denominator of each indicator, the overall quality of enrollment data affects the quality of all education indicators. Other quality assessment criteria are only appropriate for specific indicators (see section I). These include non-response rates and trends for specific items. In calculating indexes, therefore, it is convenient to classify quality assessment criteria by their applicability to a specific education indicator.

The methodology is best illustrated by example and this is shown for a fictional country in table 3.7.

Table 3.7—Calculation of Evaluation Scores for Net Enrollment and Repetition Rates for Fictional Country

	Net Enrollment		Repetition	
Evaluation Method	**score**	**out of**	**score**	**out of**
School coverage and non-response				
Coverage	2	2	2	2
School non-response	1	2	1	2
Timeliness				
No. of years from collection to dissemination	2	2	2	2
Overall content evaluation				
Enrollment trend	2	2	2	2
Cohort flow analysis	2	2	2	2
Evaluation specific to indicator				
Item non-response	N/A	N/A	0	2
Trend	N/A	N/A	0	2
Total score	9	10	9	14
Percent of total possible	90%		65%	

Evaluation Method: Population	Score	Out of
Census evaluation	0	4
Years since last census	2	2
Cohort component used	0	2
Major upheavals	0	2
Total score	2	10
Percent of total possible	20%	

The fictional country shows a score of 90% for enrollment and 65% for repetition from its administrative data. For this country a low score for population is used for illustrative purposes. Since the repetition rate is based solely on administrative data, it can be assigned an index from the administrative score. For enrollment, the score for the numerator and denominator must be averaged. The average of these two scores is 55%

3.42 ASSIGNING A PRELIMINARY DATA QUALITY INDEX

Scores for the indicators are translated to data quality index values using table 3.6. Accordingly, the repetition indicator, which has a score of 65%, is assigned an index of 3 and the net enrollment indicator, which has a score of 55%, is assigned an index of 2 (table 3.8)

Table 3.8—Preliminary Assignment of Indexes for Net Enrollment and Repetition Rates for Fictional Country

Overall score	Percentage	Index value
Net Enrollment Rate	55%	2
Repetition Rate	65%	3

3.43 TRIANGULATION

The final step is to triangulate. Namely, the value of the indicator must be assessed against all other information available, including the same indicators from other sources. If the values of the same indicator from different sources are wildly divergent, then it may be possible to reconcile these differences. Findings from such an effort may suggest that an adjustment to the index is necessary.

4. TEST ANALYSIS

In the previous section, the methodology was developed to assess data quality of selected indicators that will be provided on the data center's website. In this section, the methodology is tested using data from one or more of the data center's twenty pilot counties to ascertain its reasonableness. That is, does the methodology provide a systematic, objective means of analyzing data quality? For administrative data, Uganda will be used to test outcomes from the evaluation matrixes for reasonableness.[6] For sample survey data, again, Uganda will be used as a test case for the DHS and Ed Data. The Philippines will be used for the MICS.

4.1 Administrative and Census Data: Uganda

Uganda's indicators from administrative data and population estimates for 2003 are shown in table 4.1. Note that the net enrollment ratio has an impossible value of 101%. Clearly either the numerator is overstated or the denominator is understated, or both. This problem should manifest itself in a low quality index value. The evaluation of this indicator, as well as the other three, based on the criteria outline in Section III is presented below.

Table 4.1—Indicators, Uganda, 2003

Indicator	Reported value
Net enrollment ratio	101%
Repetition rate	14%
Completion ratio	10%
Pupil-teacher ratio	52

4.11 ADMINISTRATIVE DATA

Completeness

Evaluation criteria for completeness of the data include school coverage, school non-response, and repetition and teacher item non-response. The Ministry of Education has compiled an enumeration of schools in Uganda, which is provided in a table in the administrative database. This enumeration includes all government funded schools and all known private schools. The effort has been on-going for at least the past four years and, although it is believed that there are still some private schools not included in the database, these schools likely represent a small share of total schools. As a result of this effort, Ugandan sources believe that the coverage rate for all schools in the country in

[6] An in-depth analysis of Uganda's administrative data was conducted to provide insights into likely issues with this kind of data. This analysis is contained in Appendix 1. A summary of the results are provided here and used to construct the data quality indexes for indicators using this source.

2003 was likely in the high nineties. School non-response was measured directly from the database. For 2003 this response rate was 90%. There was no imputation or adjustment for this level of non-response.

The non-response rate for teachers was extremely low, at less than a half a percent in 2003. Conversely response rates for repetition were high. The percent of schools that reported no repetition in any grade, which peaked in 2002 at 22%, was 16% in 2003. No imputation was done for repetition non-response.

Content

Evaluation criteria for content include analyses of trends, and a check for internal consistency. Uganda's data for 2003 are extraordinarily timely, having been available since November 2003, and are currently being used for policy analysis and decision-making.

Trends in total enrollment, enrollment in the final primary grade, repetition, and the number of teachers in Uganda across the four years 2000-2003 are shown in table 4.2 These trends show remarkably large increases across time although at very different rates. Total enrollment increased by 4-8% per year; enrollment in grade 7 increased at even higher rates; the number of repeaters increased at very high and increasing rates from 3% from 2000 to 2001 to a high of 31% from 2002 to 2003; the number of teachers and enrollment in the final primary grade also increased dramatically but at a decreasing rate (table 4.2).

Table 4.2---Trends in Total Enrollment, Repetition Number of Teachers, and Enrollment in Grade 7, 2000-2003

	Enrollment		Repeaters		Teachers		Enrollment, G7	
Year	Total	% change	Total	% change	Total	% change	Total	% change
2000	6,558,576		635,004		110,366		384,368	
2001	6,899,820	5%	655,771	3%	127,038	15%	427,981	11%
2002	7,354,153	7%	774,314	18%	139,484	10%	460,109	7%
2003	7,633,314	4%	1,014,272	31%	145,703	4%	485,703	3%

Much of the increase in these indicators is a result of changes not on the ground, but in coverage and non-response rates. Estimated coverage rates increased over the four-year period, ranging from a low of 80-85% in 2000 to more than 95% by 2003. School response rates over the same period were estimated at 93% from 2000 to 2002 but declined slightly in 2003 to 90% (table 4.3).

Table 4.3--Coverage and School Response rates

Year	No. schools in the country	Annual % change in number of schools	No of Schools Surveyed	Responding Schools	Coverage rate	Response rate	Combined coverage and response rate
2000	13,793		12,480	11,576	90%	93%	84%
2001	14,201	103%	13,238	12,275	93%	93%	86%
2002	14,614	103%	14,299	13,332	98%	93%	91%
2003	14,860	102%	14,860	13,356	100%	90%	90%

When adjusted for the effect of the increasing coverage rate over the period, enrollment appears to be increasing at about the same rate as population (see Appendix I). Similarly, the increase in teachers is fully explained by two factors: the effect of the change in coverage and non-response, a data issue, combined with a deliberate government policy to increase the number of teachers in the education system. The large increases observed in repetition cannot be fully explained by school coverage and non-response, however. Even when taking this effect into account, repetition increased at a very high rate despite the current government policy to suspend repetition in schools. With regard to enrollment in the final primary grade, again, school coverage and non-response only partially explain the variations over time observed in this indicator.

The internal consistency check, which consists of an analysis of the cohort flow given by enrollment and repetition by grade, did not reveal any inconsistencies in the data (table 4.4).

Table 4.4—Calculated Promotion, Repetition and Dropout Rates by Grade, 2003

	Grade						
Item	1	2	3	4	5	6	7
Promotion rate	59%	85%	82%	78%	73%	62%	89%
Repetition rate	16%	12%	13%	13%	14%	14%	11%
Dropout rate	24%	3%	5%	9%	14%	24%	0%
Total	100%	100%	100%	100%	100%	100%	100%

4.12 CENSUS DATA

Uganda's most recent census was conducted in 2002 by the Uganda Bureau of Census (UBOS). A projection from it provides the basis for the denominator of two education indicators for 2003: the net enrollment ratio and the completion ratio. The evaluation of this population projection therefore comprises two parts, an evaluation of the census that serves as the base population and an evaluation of the post-censal estimate.

Evaluation of Census

As part of the census exercise, UBOS undertook a Post Enumeration survey (PES) to measure of the magnitude and sources of errors in the census data. The PES was carried out in a sample of 350 Enumeration Areas countrywide and took place in January 2003. The results are not yet available. The age-sex distribution of the 2002 census is also not yet available. As a stop-gap measure, UBOS applied the age distribution of the DHS conducted in 2001 to the 2002 census population (UBOS, 2003). Whereas this age distribution may be suitable for age specific population estimates until the distribution from the census is available, it is inappropriate to use it to assess the quality of the census data.

Post Censal Estimates

Evaluation criteria for post-censal population estimates include the number of years elapsed since the census, the projection method used, and information about major events that have impacted on population. The 2003 population figure, which was only one year from the last census was estimated using a simple growth rate. That is, the 2002 overall population was increased by 2.86% and the age distribution from the DHS was again applied to the total population (table 4.5).[7] As discussed in Section III, this simple mathematical model is not adequate to estimate age-specific populations during non-censal years.

[7] It is not clear what basis was used for this growth rate. The provisional report on the 2002 census indicated that the average annual growth rate from 1991 to 2002 was 3.4%.

Table 4.5-- Population by Age, Uganda, 2002-2003

Age	Population 2002	Population 2003	Percent Change
6	1,039,075	1,068,817	2.86%
7	884,990	910,320	2.86%
8	855,463	879,943	2.86%
9	718,799	739,376	2.86%
10	899,002	924,730	2.86%
11	609,877	627,335	2.86%
12	815,213	838,542	2.86%
13	699,041	719,046	2.86%
14	678,611	698,033	2.86%
15	499,489	513,792	2.86%
16	489,152	503,149	2.86%
17	428,945	441,213	2.86%
18	544,656	560,245	2.86%
19	352,331	362,408	2.86%
Total	9,514,644	9,786,949	2.86%

There has been considerable upheaval in Uganda that may be impacting on its population. During the period since the last census, the situation in the north has deteriorated. The conflict has intensified with increased incursions into northern Uganda by the rebel Lord's Resistance Army from its base of operations in southern Sudan. As a result, the eight northern districts have people moving back and forth across the Sudanese and Kenyan borders. There is also considerable internal displacement west of the Nile, for example, Gulu, Kotido, Nakapiripirit.

4.13 COMPUTATION OF INDEXES

Application of Scoring System

Scores for Uganda's administrative data are shown in diagrams 4.1 and 4.2 and summarized in table 4.6. Coverage in 2003 was estimated at 95% or more, resulting in a score for this criterion of 2 out of 2. School non-response was estimated at 90%, resulting in a score for this criterion of 1 out of 2.

Because Uganda's data was exceptionally current, it scored 2 out of 2 for timeliness. Likewise, the enrollment trend was consistent with population growth and there were no notable internal consistencies deriving from the cohort flow analysis. For both of these criteria, the data scored 2 out of 2.

A number of problems were found regarding repetition: repetition non-response was high and repetition was increasing at a very high rate despite the current government policy to suspend repetition in schools. The two evaluation criteria for this indicator—item non-response and trend analysis—score of 0 out of 2.

In contrast to repetition, teacher non-response was negligible and the trend in the level of teachers reflects government policy to increase the number of teachers in primary schools. The evaluation of teachers results in scores of 2 out of 2 for both criteria.

Finally, the trends in enrollment in the final primary grade showed a lot of unexplained movement from year to year resulting in a score for this criterion of 0 out of 2.

Overall scores for administrative data only, which are calculated as the percent of their score out of the total possible, show a score of 90% for net enrollment, 64% for repetition, 93% for the pupil-teacher ratio, and 75% for the completion ratio.

Diagram 4.1---Evaluation Matrix 1: Uganda Administrative 2003

Evaluation matrix 1: Coverage and response analyses					Score	Total
Was coverage error measured/can it be measured?	No				0	
	Yes	Coverage error is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Coverage error is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Coverage error is less than 5%			2	2
Was school non-response measured/can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2
Was repetition non-response rate measured/ can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2
Was teacher non-response rate measured/ can it be measured?	No				0	
	Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
				Yes	2	
		Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
				Yes	2	
		Non-response rate is less than 5%			2	2

Diagram 4.2---Evaluation Matrix 2: Uganda Administrative 2003

Evaluation matrix 2					Score	Total
Coverage and response analysis bypassed (if no, skip this criterion)					0	2
Content	Are there at least four years of sequential data for net enrollment?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years of sequential data for repeaters?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years of sequential data for teachers?	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are there at least four years data for enrollment in the final primary grade	No			0	
		Yes	Unexplained percent change in level varies from one year to the next by	More than 3%	0	
				Between 1% and 3%	1	
				Less than 1%	2	2
	Are enrollment and repetition by grade available for two consecutive years?	No			0	
		Yes	Are all residual dropout rates positive?	No	1	
				Yes	2	2
Timeliness	Number of years elapsed from the time the data was collected until it was disseminated			More than two years	0	
				Two years	1	
				One year	2	2

Table 4.6--Scores Obtained from Evaluation of Uganda's Administrative Data, 2003

	Net Enrollment		Repetition		Pupil-Teacher		Completion	
Evaluation Criterion	**score**	**out of**	**score**	**out of**	**score**	**out of**	**score**	**out of**
School coverage and non-response								
Coverage	2	2	2	2	2	2	2	2
School non-response	1	2	1	2	1	2	1	2
Timeliness								
No. of years since last available Estimate	2	2	2	2	2	2	2	2
Overall content evaluation								
Enrollment trend	2	2	2	2	2	2	2	2
Cohort flow analysis	2	2	2	2	2	2	2	2
Evaluation specific to indicator								
Item non-response	N/A	N/A	0	2	2	2	N/A	N/A
Trend	N/A	N/A	0	2	2	2	0	2
Total score	9	10	9	14	13	14	9	12
Percent of total possible	90%		64%		93%		75%	

Uganda's population estimates score very low for data quality as indicated in diagram 4.3 and 4.4. Uganda has not yet provided it's own analysis of the quality of the census or adjustment factors. Therefore, evaluation matrix 1 is bypassed. It has not provided an age-sex distribution from the census. Therefore there is no basis on which to conduct an in-house evaluation. Lacking an evaluation, the census data scores a value of 0 out of 6 for this evaluation method.

Only one year has elapsed since the census, giving Uganda's projection a score of 2 out of 2 for this criterion. However, it was based on a simple mathematical calculation rather than a cohort component method, resulting in a score of 0 out of 2 for this evaluation criterion. Finally, internal and external unrest continues to cause considerable movement across Uganda's borders, rendering it more difficult to make accurate population projection, thus the final evaluation criterion—major upheavals—receives a score of 0 out of 2 (table 4.7).

Diagram 4.3—Evaluation Matrix for Census Data and Post Census Estimates: Census and Population Estimates, 2003

Evaluation Matrix 1c—Census evaluation conducted in-house				Score	Total
Is age-sex distribution available?	No ——— go to evaluation matrix 2			0	4
	Yes	Is graphical analysis of age-sex distribution consistent with expectations?	Yes	0	
			There are some irregularities	1	
			There are distinct preferences for particular digits	2	2
		UN age-sex index	Over 40	0	
			Between 20 and 40	1	
			Under 20	2	2

Diagram 4.3 cont.—Evaluation Matrixes for Census Data and Post Census Estimates: Census and Population Estimates, 2003

Evaluation Matrix 2-- Post-census estimates		Score	Total
Score from evaluation matrix 1		0	4
Number of years since last census	1-4	2	
	5-8	1	
	More than 8	0	2
Are post census estimates based vital registers or on cohort component method?	Yes	2	
	No	0	2
Since the last census, have there been any upheavals causing unusual migration or death rates?	No	2	
	Events of a short term nature may have effected migration rates, but the effect is expected to be small	1	
	There have been major upheavals known to have impacted population	0	2

Table 4.7--Evaluation of Uganda's Population Projection, 2003

Evaluation Method	Score	Out of
Census evaluation	0	4
Years since last census	2	2
Cohort component used	0	2
Major upheavals	0	2
Total score	2	10
Percent of total possible	20%	

Assigning a Preliminary Index

The repetition rate and the pupil-teacher ratio are calculated only from administrative data. Their scores were 64% and 93% respectively. Using the breakdown in section III to translate a percentage into and index value, repetition receives an index value of 3 and the pupil teacher ratio receives an index value of 4.

The index for the net enrollment rate and the completion ratio is calculated as the average of the scores for the numerator and denominator (table 4.8). The very low score in the denominator results in low overall scores for net enrollment and completion ratios.

Table 4.8—Calculation of Preliminary Index for Net Enrollment and Completion Ratio for Uganda, 2003

Indicator	Numerator	Denominator	Average of Numerator and Denominator	Index
Net enrollment ratio	90%	20%	55%	2
Completion ratio	75%	20%	48%	2

Returning to the table of indicators shown in table 4.1, the indexes have been added in table 4.9. Are they reasonable? The pupil-teacher ratio gets the highest index of 4; the response rate for teachers was very high, at 99%; the data are timely; and trends were consistent with national policy. This ratio, therefore, appears to be very reliable and a score of 4 is appropriate.

Table 4.9—Indicators and their Preliminary Data Quality Index, Uganda, 2003

Indicator	Reported value	Index
Net enrollment ratio	101%	2
Repetition rate	14%	3
Completion ratio	58%	2
Pupil-teacher ratio	52	4

The repetition ratio was assigned an index value of 3. There was a high non-response rate for this indicator and the trend shows unexplained increases in repetition. On the other hand, the data are timely, the cohort flow analysis shows a reasonable progression of pupils through the system with the repetition and enrollment data. An index of 3 indicates that the indicator is fairly reliable, but that trend analysis and cross-country comparisons require caution. This value seems reasonable.

The net enrollment and completion ratios both received an index score of 2. The numerators had fairly high scores. The denominators did not. The low score in the

denominators had the effect of pulling the overall score down from what would have been the highest quality index to a value of only 2.

4.2 DHS: Uganda

A Demographic and Health Survey and Ed Data Survey were conducted in Uganda in 2001. The DHS enumerated members of a household. Its target population was women in the reproductive age. The Ed Data survey interviewed a parent or guardian about each of his or her children enumerated in the DHS regarding their education and factors affecting it.

4.21 NON-RESPONSE

Child Non-Response

Since parents or guardians responded for their children and children for whom they are responsible, the child response rate is a product of the household response rate, the parent/guardian response rate, and the eligible child response rate (Macro, 2002, p8). The overall response rate for children was 96%.

Item Non-Response

Households were considered eligible for the Ed Data survey if the head of the household was more than 19 years of age. For all children in such households aged 5-18 identified in the DHS, parents or guardians were asked whether they attend school regularly and the grade in which they were currently enrolled. Grades ranged from pre-school through primary, secondary and tertiary. Non-response rates for these two questions were infinitesimal as shown in table 4.10.

Table 4.10--Attendance and Grade Non-Response Rates

Item	No Of Non-Responses	Number of Relevant Cases	Non-Response Rate
Is (NAME) currently attending school?	4	11,614	0.0%
What class/year is (NAME) currently attending?	1	8,334	0.0%

4.22 CONTENT

Age structure

The distribution of household populations by single ages is shown in figure 4.1. It is readily apparent from this figure that there is some age misreporting, however, there does not appear to be any single-digit preference.

Figure 4.1--Single Age-Sex Pyramid, Uganda 2001 DHS
Percent of Total Population

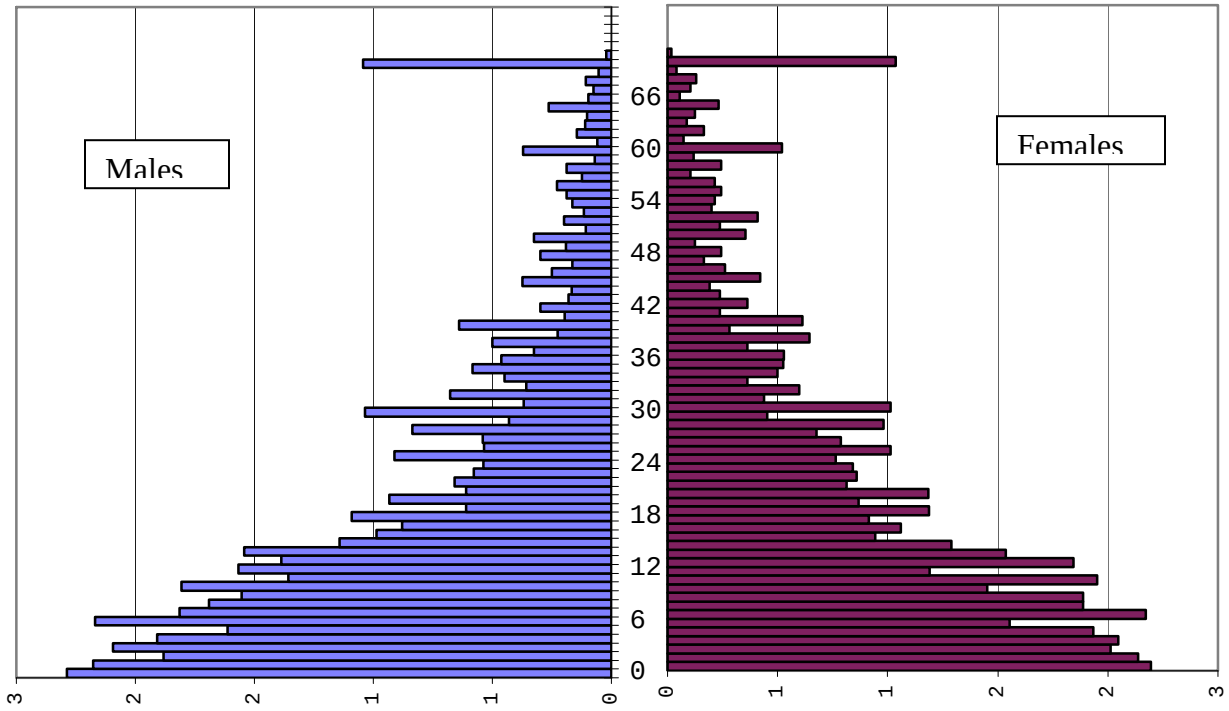
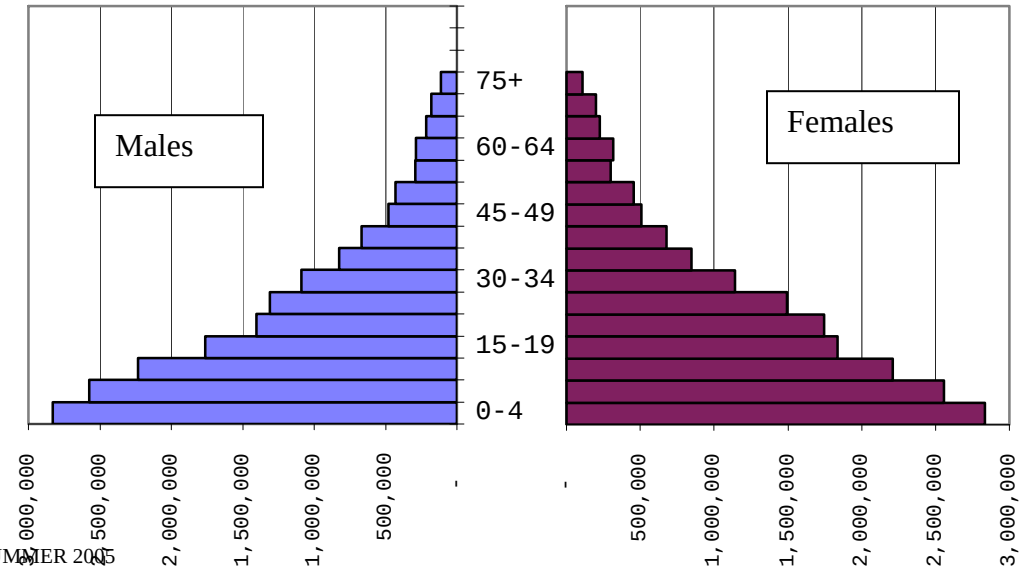


Figure 4.2--Population Distribution by Five-Year Age Groups by Sex



The distribution of the population by five-year age groups by sex is shown in figure 4.2. This figure is also provided in the DHS report. When aggregated across five-year age groups, variations across ages are smoother. Since the indicator being measured is for the age group 6-12 years old, the variations in single age groups is not an issue.

The age distribution of Uganda's population shows a young age structure, with over half the population under 15. The DHS report notes that this type of age structure has a built-in momentum for growth of a country's population. This is consistent with findings from the recent population census of a growth rate of 3.4%. The DHS notes that there is an unexpected bulge in the proportion of women in the 50-54 age group. According to the report, this anomaly is likely due to women age 45-49 being deliberately reported in the 50-54 age category to reduce the workload of the interviewer. Since women in this age category are well above the school age population, it is not a concern here.

In summary, although the age-sex pyramid shows wide variations by single age, these anomalies become minimal when aggregated across five-year age groups. The overall age structure is consistent with expectations.

Enrollment by Age by Grade

Enrollment by age by grade is shown in table 4.11. This cross-tab shows some children of very young ages in some of the grades. In addition, some of the grades were outside of the range for primary school. Total enrollment in these cells was small, however, at less that half a percent.

Table 4.11--Enrollment in Primary by Grade, Uganda Ed Data Survey, 2001

	Grade												Total in Unlikely
Age	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**12**	**13**	**14**	**16**	**Total**	**Cells**
5	183	*7*	*1*	*1*	*1*							193	10
6	565	108	*4*	*1*								678	5
7	535	276	58	*2*	*1*		*4*					876	7
8	282	420	251	39	*4*							996	4
9	126	258	292	137	25	*2*						840	2
10	68	181	368	235	134	18						1,004	-
11	26	83	157	213	132	60	9					681	-
12	18	89	179	233	232	106	66	*1*	*1*	*1*	*1*	927	4
13	2	28	93	175	191	161	93		*1*			744	1
14	3	17	42	114	151	164	110					601	0
15	2	5	15	40	104	124	99					389	0
16	1	2	11	17	44	64	88					227	0
17		1	2	4	10	30	51					98	0
18		1	1	4	10	18	33					67	0
19					1	4	8					13	0
Percent of total enrollment in unlikely cells: 0.4												8,334	33

Sampling Error

The Ed Data survey provided estimates of errors for selected ratios. The enrollment ratio for which they reported the sampling error was the total of all children age 6-12 enrolled in school divided by the sum of all children age 6-12. Children enrolled in school included those enrolled at all levels—preprimary, primary and post-primary. As enrollment ratios are normally reported for each education level separately, it is not clear why the standard error of this particular ratio was reported.[8] This ratio is 91.2%, compared to the primary school net attendance ratio of 87.1%.

4.23 CONSTRUCTION OF DATA QUALITY INDEX

The evaluation matrix for Uganda's DHS plus Ed Data is shown in diagram 4.5 and table 4.12. This survey received full points on all criteria with the exception of the standard error of the estimate. Again, it is not clear why the report did not give the standard error for the primary net enrollment ratio, as this was one of the principle statistics used in the main body of the report, as well as one of the most commonly cited education statistics. The total score was 10 out of 12, or 83%, giving it an index of 4.

[8] Nor is it documented that this is the ratio that they are reporting. The main body of the report gives the more traditional ratio of attendance in *primary* school of 6-12 year-olds divided by the population of 6-12 year-olds. It is only through downloading the data that the exact computation can be ascertained.

Diagram 4.5—Evaluation Matrix for DHS/EdData, Uganda 2001

Evaluation Matrix						Score	Total
Unit and item non-response	Was household/ person non-response measured	No				0	
		Yes	Non-response rate is greater than 10%	Was adjustment made?	No	0	
					Yes	2	
			Non-response rate is 5-10%	Was adjustment made?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
	Was non-response rate measured for attendance?	No				0	
		Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
					Yes	2	
			Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
	Was non-response rate measured for grade?	No				0	
		Yes	Non-response rate is greater than 10%	Was adjustment made?	No	0	
					Yes	2	
			Non-response rate is 5-10%	Was adjustment made?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
Content	Is age sex distribution available?	No				0	
		Yes	Is age-sex pyramid consistent with expectations?		No	0	
					Yes	2	2
	Is attendance/ enrollment by grade by age available?	No				0	
		Yes	Percent of pupils in unlikely or impossible grade-age cell less than 3%		No	1	
					Yes	2	2
Sampling error	Was sampling error provided?	No				0	
		Yes	95% confidence interval less than or equal to ± 2%			1	
			95% confidence interval greater than to ± 2%			2	2

Table 4.12-- Scores Obtained from Evaluation of the DHS/EdData, Uganda, 2001

Criterion	Score	Out of
Child Non-Response	2	2
Non-Response: Attendance	2	2
Non-Response: Grade	2	2
Age Structure	2	2
Enrollment by Grade	2	2
Standard Error	0	2
Total	10	12
Percent of total	83%	

4.3 Triangulation: The Final Evaluation Criterion

Table 4.13 shows the indicators from the two sources—administrative/population estimates and the DHS. The net enrollment rate reported by administrative sources was 101% compared to a DHS net attendance ratio of 87%. If the DHS showed net enrollment rates close to 100%, the rate provided by the administrative data although slightly greater than 100% might be explained by small errors in the numerator or denominator. However, the difference between the estimates from the two sources is quite large.

Table 4.13—Indicators and Preliminary Data Quality Indexes from Administrative Sources and the DHS, Uganda, 2003

Indicator	Administrative Reported value	Data Quality Index Value	DHS Reported value	Data Quality Index Value
Net enrollment (attendance) ratio	101%	2	87%	4
Repetition rate	14%	3	-	-
Completion ratio	58%	2	-	-
Pupil-teacher ratio	52	4	-	-

Triangulation has already been applied to some extent through out the evaluation with the following results. The administrative/population data has an index value of 2 for the following reasons:

- The numerator, which is derived from administrative data, scored very well in terms of coverage, timeliness, trend analysis and cohort flow analysis.
- However, there was a 10% school non-response rate for which no adjustment was made.
- The population estimate scored poorly for a plethora of reasons.

That there was a 10% school non-response rate for which no adjustment was made means that the numerator is likely understated. However, the ratio adjusted for this non-response rate would be even greater than 101%. Meanwhile, there is a very good chance that population census was under-enumerated. Post-enumeration surveys generally show that population censuses usually are under-enumeration on the magnitude of 1-3 percent. This level of under-enumeration, however, is not enough to reduce the net enrollment rate much below 100% and does not explain the large difference between it and that provided by the DHS. All of this suggests that there is still an issue with the administrative data that was not revealed in the evaluation thus far.

The first and most obvious explanation for the difference between the two sources is that the DHS uses attendance in its numerator and the administrative survey uses

enrollment. To the extent that pupils enroll at the beginning of the year, but do not actually attend, the enrollment ratio would be higher than the attendance ratio.

Another critical factor in estimating net enrollment is age. The DHS obtained information about the age of the children in the household from the parent and then collected information about school attendance. Age reported by schools also ultimately comes from parents. However, where records are incomplete or the school administrator was hasty in completing the questionnaire, there is a large potential for misreporting.

The average age by grade calculated from the DHS is compared to an estimate of average age by grade derived from the administrative data shown in tables 4.14. Unfortunately the age distribution in the administrative questionnaire collapses all pupils over age 14 into one age category. Therefore the average age calculated from this source tends to be understated in the higher grades. However it does show that even in the lower grades, average age is considerably higher in the DHS than in the administrative data.

Table 4.14—Average Age by Grade from Administrative Data and DHS, Uganda 2003

Grade	Age	
	Admin Data	DHS
1	6.7	7.0
2	8.2	8.7
3	9.6	10.1
4	10.7	11.5
5	11.7	12.6
6	11.7	13.7
7	12.5	14.6

Another way to evaluate the differences in ages reported by the two sources is to compare the percent under- and over-age pupils. The comparison, presented in table 4.15, shows that overage pupils represent a much larger share of the total student population shown by the DHS than is shown by the administrative data.

Table 4.15—Percent of Pupil Population that is Under- and Over-Age, Uganda 2003

Student Population	Admin Data	DHS
Underage	4%	2%
Overage	17%	26%

Taking all factors together, it seems extremely likely that age misreporting occurred in the administrative data. If so, then the score obtained for the numerator of the net enrollment ratio must be downgraded and the data quality index for the indicator from administrative data should be reduced to a value of 1 (table 4.16). Note that an in depth

data quality analysis of Uganda's data, confirms that age misreporting is a significant issue.

Table 4.16—Indicators and Final Data Quality Indexes from Administrative Sources and the DHS, Uganda, 2003

Indicator	Administrative Reported value	Data Quality Index Value	DHS Reported value	Data Quality Index Value
Net enrollment (attendance) ratio	101%	1	87%	4
Repetition rate	14%	3	-	-
Completion ratio	58%	2	-	-
Pupil-teacher ratio	52	4	-	-

4.4 MICS: Philippines

The Multiple Indicator Cluster Survey (MICS) was developed by UNICEF to provide countries with a survey methodology using probability sampling to assess progress on all goals of the World Summit for Children at the end of the decade. A MICS survey was conducted in Philippines in 1999. This survey enumerated members of households in the sample. Data was collected for all children enumerated in the household aged 5-17.

4.41 NON-RESPONSE

The household non-response rate for the MICS survey in the Philippines was 97.8%.[9] For all children enumerated in households that were aged 5-17 parents or guardians were asked whether they had ever attended school, if they were currently attending school, and whether they attended school at any time during the current year. If children were currently attending school or had ever attended school during the current year, they were asked what grade they attended. Non-response rates for these four questions were infinitesimal as shown in table 4.17.

Table 4.17--Attendance and Grade Non-Response Rates

Item	No Of Non-Responses	Number of Relevant Cases	Non-Response Rate
Has (*NAME*) ever attended school?	29	11,314	0.3%
Is (*NAME*) currently attending school?	3	9,829	0.0%
During the current school year, did (*NAME*) attend school at any time?	3	1031	0.0%
Which level and grade is/was (*NAME*) attending?	0	8,903	0.0%

4.42 CONTENT

Age structure

The distribution of household populations by single ages, provided in figure 4.3, shows that there appears some age misreporting with a slight preference for digits ending in 0. this age reporting preference appears to be slightly more apparent in the older age categories. Since the indicator being measured is for the age group 6-12 years old, the variations in single age groups is not an issue.

[9] This non-response rate was calculated from the frequency distribution of the field called "Result of Household Interview". The report provided on the website cites a household non-response rate of 98.4%. Although it is not clear why there is this difference, it is not of concern with such low non-response rates.

The distribution of the population by five-year age groups by sex is shown in figure 4.4. When aggregated across five-year age groups, variations across ages are very smooth.

Figure 4.3--Single Age-Sex Pyramid, Philippines 1999 MICS
Percent of Total Population

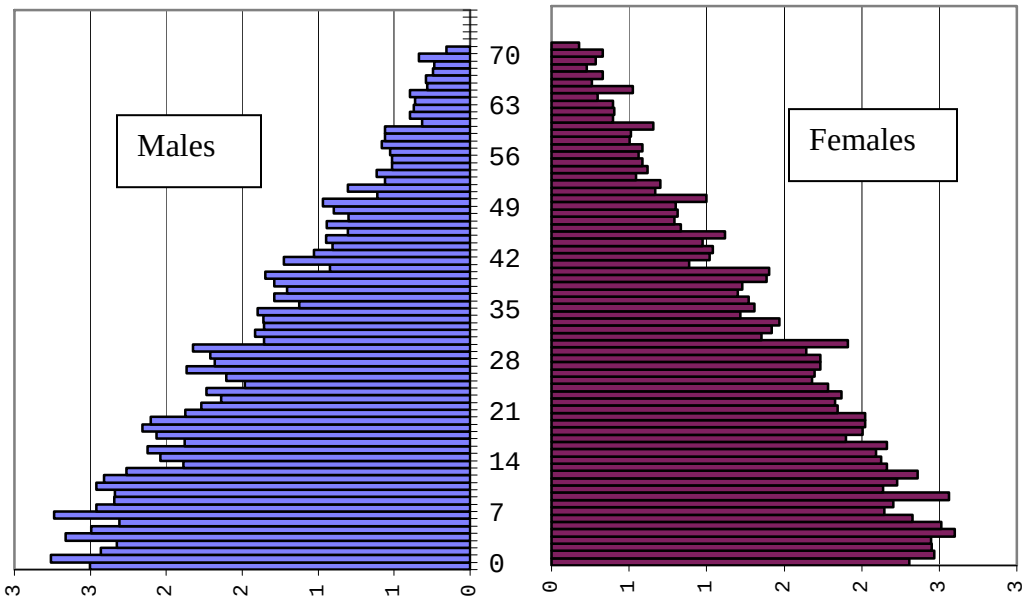
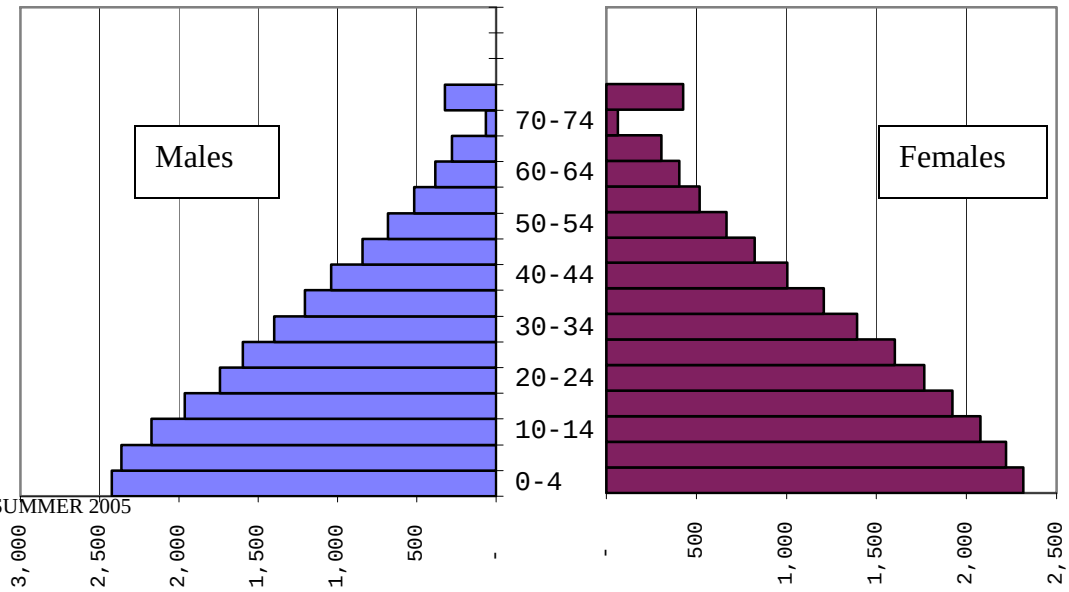


Figure 4.4--Population Distribution by Five-Year Age Groups by Sex, Philippines 1999 MICS



In general, the population distribution of the Philippines' shows a pyramid with a wide base and narrow top, a pattern that is typical of populations with high fertility rates, and the overall age structure is consistent with expectations.

Enrollment by Grade

Enrollment by age and grade is shown in table 4.18. This cross-tab shows very few children in unlikely cells. It is not clear what is meant by the grade level labeled "Graduate".

Table 4.18--Enrollment in Primary by Grade, Philippines MICS, 1999

	Grade								
Age	**1**	**2**	**3**	**4**	**5**	**6**	**Graduate**	**Total**	**Total in Unlikely Cells**
5	14	*1*						15	1
6	291	5						296	5
7	532	254	15	*2*				803	2
8	135	461	228	16	*1*			841	1
9	45	160	446	222	13	*2*		888	2
10	25	57	140	384	179	10		795	0
11	10	27	74	129	469	98		807	0
12	8	19	34	74	150	467	*1*	753	1
13		4	23	26	41	143		237	0
14		2	3	13	24	44	*1*	87	1
15		1	2	11	9	22		45	0
16			2	4	6	10		22	0
17				3	7	5		15	0
	1,060	991	967	884	899	801	2	5,604	13

Sampling Error

The sampling error of the net enrollment rate was not provided.

4.43 CONSTRUCTION OF DATA QUALITY INDEX

The evaluation matrix for the Philippines' MICS is shown in diagram 4.6 and scores are summarized in table . This survey received full points on all criteria with the exception of the standard error of the estimate, giving it an index 4.

Diagram 4.6—Evaluation Matrix for MIC, Philippines 1999

Evaluation Matrix						Score	Total
Unit and item non-response	Was household/ person non-response measured	No				0	
		Yes	Non-response rate is greater than 10%	Was adjustment made?	No	0	
					Yes	2	
			Non-response rate is 5-10%	Was adjustment made?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
	Was non-response rate measured for attendance?	No				0	
		Yes	Non-response rate is greater than 10%	Was imputation / adjustment done?	No	0	
					Yes	2	
			Non-response rate is between 5-10%	Was imputation / adjustment done?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
	Was non-response rate measured for grade?	No				0	
		Yes	Non-response rate is greater than 10%	Was adjustment made?	No	0	
					Yes	2	
			Non-response rate is 5-10%	Was adjustment made?	No	1	
					Yes	2	
			Non-response rate is less than 5%			2	2
Content	Is age sex distribution available?	No				0	
		Yes	Is age-sex pyramid consistent with expectations?		No	0	
					Yes	2	2
	Is attendance/ enrollment by grade by age available?	No				0	
		Yes	Percent of pupils in unlikely or impossible grade-age cell less than 3%		No	1	
					Yes	2	2
Sampling error	Was sampling error provided?	No				0	
		Yes	95% confidence interval less than or equal to ± 2%			1	
			95% confidence interval greater than to ± 2%			2	2

Table 4.19-- Scores Obtained from Evaluation of the MICS, Philippines, 1999

Criterion	Score	Out of
Child Non-Response	2	2
Non-Response: Attendance	2	2
Non-Response: Grade	2	2
Age Structure	2	2
Enrollment by Grade	2	2
Standard Error	0	2
Total	10	12
Percent of total	83%	