THE OPTIMAL REFERENCE GUIDE

The Data Quality Imperative

Data Quality Series—Part 1

Extraordinary insight™ into today's education information topics

By Glynn D. Ligon, Ph.D., ESP Solutions Group, Inc.

Table of Contents

Foreword iii

Introduction ..... - 1 -

The Imperative ..... - 2 -

Perspectives of Practitioners – How professionals who manage data view data quality ..... - 5 -

Background ..... - 5 -

The Pursuit of a Definition of Data Quality ..... - 7 -

A Hierarchy of Data Quality – Getting to data-driven decision making ..... - 10 -

Bad Data ..... - 10 -

-1.1 Invalid ..... - 10 -

None ..... - 11 -

0.0 Unavailable ..... - 11 -

Available ..... - 12 -

1.1 Inconsistent Forms of Measurement ..... - 12 -

1.2 Data Collected by Some at Some Times ..... - 12 -

1.3 Data Combined, Aggregated, Analyzed, Summarized ..... - 13 -

Official ..... - 15 -

2.1 Periodicity Established for Collection and Reporting ..... - 15 -

2.2 Official Designation of Data for Decision Making ..... - 15 -

2.3 Accuracy Required for Use in Decision Making ..... - 16 -

Valid ..... - 16 -

3.1 Accurate Data Consistent with Definitions ..... - 16 -

3.2 Reliable Data Independent of the Collector ..... - 17 -

3.3 Valid Data Consistent with the Construct Being Measured ..... - 18 -

Quality ..... - 18 -

4.1 Comparable Data: Interpretable Beyond the Local Context ..... - 18 -

4.2 Data-Based Decisions Made with Confidence ..... - 19 -

Steps for Ensuring Data Quality ..... - 20 -

Conclusion ..... - 20 -

Attachment A ..... - 21 -

Attachment B ..... - 23 -
Data Quality Boot Camp – Understanding the principles of data quality ..... - 23 -
Principles of Data Quality ..... - 23 -
Attachment B: Process Illustration of Data Quality 27

Foreword

Data quality is akin to writing down your own personal PIN number in order to remember it.

This may be the best nonsports analogy up with I've ever come. Think of a PIN as a datum, a fact, a single piece of information that is crucial to answering the most important question of the moment—are you really who you claim to be? That determined, then the system releases to you information and the capacity to take action using that information. Now isn't that what our education information systems are really all about?

Let's return to our original sentence and admit that it's an egregious pleonasm.

> **Ple-o-nasm** *n* to use more words than necessary to denote mere sense; antonym: oxymoron

Clearly, "personal PIN number" is pleonastic to a fault, but what about "data quality"? Shouldn't all data be quality data before we allow them into our longitudinal data systems? Shouldn't our longitudinal data systems be safe havens for quality data? Oh my, safe haven is a pleonasm, isn't it? Users of our longitudinal data systems don't expect to pay extra for data quality, they expect quality data as a free gift—oops, another pleonasm.

Seriously now, unless you failed to see any humor in any of this, data quality should be a pleonasm if we design and manage our information systems perfectly. We should just say, "data" and say it with confidence. Data—a word that symbolizes quality. (BTW, "high quality" is considered a pleonasm to most purists; so, this paper follows that style of usage.)

Back to our PINs. A PIN is clearly defined, e.g., four numerals because the spouse of the inventor couldn't remember six. A PIN must be entered perfectly. A PIN links a specified individual with a specified identity management system for specified authorized actions during a certain period of time. If a PIN doesn't work, guess who fixes it? The user who set it in the first place. In fact, that user feels like the PIN's intrinsic, unstated quality protects the rest of the data.

ESP Insight

Our position has always been that data quality is dependent upon the initial entry of the data and the careful shepherding of the data throughout the system, including reporting.

I've never heard a DBA refer to "PIN quality." That's because the data providers are given the business rules and are held totally accountable for providing correct PINs. Once a PIN gets into the information system as a datum, it's "correct." = "quality." (I always thought it interesting that individuals' PINs do not need to be unique. In fact, everyone could have the same PIN. There are only 10,000 of them in most systems—the notable exception being in Switzerland, where they use 6 numerals.)

Data quality should never be an oxymoron.

ESP was contracted to help create a Data Quality Manual for the Office of the Chief Information Officer of the U.S. Department of Education. Another contractor was also expected to help on the development of this manual. In fact, the other contractor believed that this was their unique area of expertise, and that they alone should develop the manual. When we pressed them for the areas they would cover in the manual, they described business rules and other data cleansing techniques. We convinced our client that this was only a small part of data quality. Our expanded view of data quality is highlighted in this paper.

Our position has always been that data quality is dependent upon the initial entry of the data and the careful shepherding of the data throughout the entire data system, including reporting. This paper makes the case for data quality as no one else typically does – basically from the initial transaction/behavior being measured/reported to proper use of the ultimate information.

Introduction

The secret to quality data is simple—get them right from the beginning.

Then keeping them right is a matter of effective data management processes.

What's the tell-tale sign that an agency does not have control over the quality of its data? If an education agency is spending time cleaning data, then the processes are wrong. Cleaning data is one of the least desirable tasks for an agency. The act of cleaning data means that someone earlier in the process did something wrong.

ESP Insight

The secret to data quality is simple—get them right from the beginning.

One of the biggest mistakes that an education agency makes when a new information system project is implemented is letting bad data get into the system. Bad data must be met at the door and turned away for the provider to correct. By the way, that door needs to be as early in the process as possible.

This paper presents the clearest look into the dynamics of data quality yet developed by ESP's professionals. The reason—we've been helping education agencies improve their data quality since 1993. Before that, some of us were the ones sending in the data.

ESP Insight

The tell-tale sign that an agency does not have control over data quality is the amount of time spent cleaning data.

A simple test for the right attitude about data quality is how an agency reads these words.

CLEAN DATA

If this is read clean as an action verb—that's trouble. If it's read as a statement of pride—there's hope.

The Imperative

Much talk buzzes around data quality. This Optimal Reference Guide (ORG) describes quality data quality. That would be data quality of the highest order. That quintessential level of data quality is defined as:

> **Data quality is more than accuracy and reliability. High levels of data quality are achieved when information is valid for the use to which it is applied, and decision makers have confidence in the data and rely upon them.**

ESP Insight

Data quality matters whether we are rating, funding, or selecting schools.

Samples of higher quality education data:

1. An official transcript certified by a high school registrar
2. A teacher's grade book for the end of a grading period—a week later
3. Teacher's certificate or license showing areas of certification or endorsement
4. Directory of sign-ons and passwords for a secure application
5. Payroll data—a month later
6. GIS file of addresses for enrolled students
7. Lunch eligibility status and meals served
8. Won-lost records of athletic teams in official events
9. School AYP status—after appeals

Samples of data that fall short of dependable quality:

1. Discipline data
2. Dollar amount of scholarships earned by graduating seniors
3. Student mobility rate
4. Student dropout rate
5. Instructional expenditures
6. Persistently dangerous schools
7. Hits on a school's website—what's a hit mean anyway?
8. Percent of high school graduates attending college—out of state
9. Number of ADHD students enrolled in kindergarten

When we rate schools and fund schools, data quality matters. When we describe schools out of curiosity (e.g., mobility rate, hits on a web site), data quality makes comparisons valid. When we select schools for our own kids to attend, softer data like anecdotes and opinions of trusted friends often trump the statistics—quality or otherwise. Whatever the purpose, we all want the best data possible.

ESP Insight

The "Four Great Truths" about data quality have held up after more than a decade of work with education agencies on quality issues.

As detailed later on, there are four great truths about data quality:

> **Data quality is highest when…**
>
> 1. The data providers know what's expected.
> 2. The data providers use the data themselves for their own work.
> 3. Everyone, everywhere checks the data.
> 4. The data are available and used.

How do these match with the infrastructure components ESP has identified thorough our experiences? See Figure 2.

Figure 2: The Truth about Infrastructure Components

Truths about Data Quality	Infrastructure Components						
	Information Systems Architecture	Infra-structure	Collec-tions	Data Stores	Decision Support System	Portal	User Support
1. The data providers know what's expected.	**X**	X	X	X	X	X	**X**
2. The data providers use the data them-selves for their own work.	**X**	X	X	X	**X**	X	X
3. Everyone, every-where checks the data.	**X**	X	**X**	X	X	X	X
4. The data are available and used.	**X**	**X**	X	**X**	X	**X**	X

ESP Insight

Education agencies need all seven D3M infrastructure components to produce quality data from a comprehensive information system.

Every component is supportive of all four great truths about data quality. The larger, green Xs indicate the primary contribution of each component. Information Systems Architecture, with its emphasis on Data Governance, is the common denominator for data quality across the board.

1. **Information Systems Architecture**—the metadata, hardware, software, and network standards, policies, governance, and requirements by which all technology systems are built and managed

2. **Infrastructure**—the physical hardware, software, network, and human resources required to support the technology systems

3. **Collections**—the mechanisms for gathering data

4. **Data Stores**—the centralized locations where data are located, managed, and accessed; includes a comprehensive data model

5. **Data-Driven Decision Support System**—the way the data are provided to users for decision making, e.g., reports, queries, data files, etc.

6. **Portal**—the system that authenticates and authorizes all users to provide appropriate access and security to all information

7. **User Support**—the system that trains, helps, and guides users to ensure efficient and proper use of the information

Perspectives of Practitioners – How professionals who manage data view data quality

The following has been revised and enhanced since first being published as Data Quality: Earning the Confidence of Decision Makers, a paper presented at the annual meeting of the American Educational Research Association, April, 1996.

> **Data quality is more than accuracy and reliability. High levels of data quality are achieved when information is valid for the use to which it is applied, and when decision makers have confidence in the data and rely upon them.**

Professionals responsible for education data have long sought to provide timely and useful information to decision makers. Regardless of the evaluation model, research design, or statistical methodology employed, informing the decision making process with quality, reliable data is the basic goal. In the publications describing quality related to general information systems, the concept is narrowly interpreted to mean accurately and reliably processed data. This section ties together the foundations of data quality from the formal information systems literature with the practical aspects of data quality in the arena of public education decision making. A hierarchy of data quality has been developed to assist both the understanding of quality and the requirements for achieving quality. The hierarchy ranges from the availability of dysfunctional, bad data to the quality level of data-based decisions made with confidence.

ESP Insight

What are not universally sought are federally imposed standards for data and information systems.

Background

Data quality is essential to successful research, evaluation, and statistical efforts in public schools. As statewide accountability systems that rely upon Big Data grow, concern follows about the data quality within those emerging state-level databases. As states and the federal government expand and institutionalize their P20W statewide longitudinal data systems to make information available electronically to everyone, questions are raised about the quality of the data collected and stored.

There is broad support for voluntary standards which states and local school districts can adopt (e.g., CEDS). What is needed first is a way to know when quality data are available and when caution should be exercised. All this must be accomplished within the context of the ever-changing world of information technology.

Decision makers at all levels are relying upon data to inform, justify, and defend their positions on important issues. What are the key criteria on which to

determine data quality? Is there a logical sequence to the processes for ensuring quality in information systems?

The concern for data quality is somewhat different than the slowly emerging interest in education data that has grown for decades. The concern for data quality is a sign of maturity in the field, an increasing sophistication by the audiences who use education data. In other words, first we asked "Are our students learning?" Then we had to ask "What are the education indicators that we should be monitoring?" Finally, we are asking "Now that we have some indicators, do we trust them?"

ESP Insight

The concern for data quality is a sign of maturity in the field, an increasing sophistication by the audiences who use education data.

An easy point in time to mark is the release of the "Nation at Risk" report. Much reform in education followed, including expansion of accountability systems within states. The search heated up for the true, reliable indicators of quality in education. Another major event was the passage of the 1988 Hawkins Stafford Education Amendments that called for improving the quality of the nation's education data. From that legislation, the National Forum for Education Statistics was begun, and from that group has followed a continuing focus on data quality issues. The Forum, sponsored by the National Center for Education Statistics, which is part of the Institute for Education Sciences, is made up of state education agency representatives and local education agency staff.

Then in 2001, everything was ratcheted up several notches with the passage of the No Child Left Behind Act. SEAs suddenly began taking the data challenges presented by accountability mandates very seriously.

There are multiple perspectives, each with its own reality of data quality. These are:

- Decision Makers (parents, teachers, counselors, principals, school board members, legislators, governors)
- Program Managers (directors, supervisors)
- General Audiences (news media, taxpayers, businesses)
- Data Collectors and Providers (clerks, teachers, counselors, program managers)
- Analysts (evaluators, researchers)

Individuals may occupy more than one of these groups simultaneously.

At the risk of over simplifying, the primary perspective of each group may be described as:

Decision Makers:
"Do I have confidence in the data and trust in the person providing them?"

Program Managers:
"Do the data fairly represent what we have accomplished?"

General Audiences:

ESP Insight

In the end, the audiences (e.g., program managers, decision makers, and general audiences) give the ultimate judgment of quality when they use, ignore, or disregard the data.

"Did I learn something that appears to be true and useful, or at least interesting?"

Data Collectors and Providers:
"Did the data get collected and reported completely and in a timely manner?"

Evaluators, Researchers, Analysts:
"Are the data adequate to support the analyses, results, and interpretations from them?"

The burden for data quality traditionally falls to the data collectors and providers. Who else would be in a better position to monitor and judge data quality? However, in the end, the audiences (e.g., program managers, decision makers, and general audiences) give the ultimate judgment of quality when they use, ignore, or disregard the data. Our conclusion? *The highest level of data quality is achieved when information is valid for the use to which it is applied and when decision makers have confidence in the data and rely upon them.*

The Pursuit of a Definition of Data Quality

Years ago, Robert Friedman, formerly the director of the Florida Information Resource Network (FIRN), Arkansas's statewide network, and the California Student Information System (CSIS), called me and asked for references related to data quality. The issue had arisen as the new statewide education information system for Arkansas was being developed. There were few references available, none satisfactory. I began documenting anecdotes, experiences, and insights provided by individuals within the education research, evaluation, and information systems areas to search for "truths." Three years after Friedman's inquiry, I responded with the following insights.

Several ideas were consistently referenced by individuals concerned with data quality.

ESP Insight

A key element frequently cited as basic for achieving quality is the reliance upon and use of the data by the persons responsible for collecting and reporting them. This may be the most important truism in this paper.

1. Accuracy
Technical staff mention reliability and accuracy. This is consistent with the published literature in the information systems area. Accuracy, accuracy, accuracy—defined as do exactly what we are told, over and over. Not all information specialists limit themselves to the mechanical aspects of accuracy; however, because they may not be content or process specialists in the areas they serve, their focus is rightfully on delivering exactly what was requested. After all, that is what the computer does for them.

Quality data in, quality data out.

2. Validity

However, programmatic staff point out that data must be consistent with the construct being described (i.e., validity). If their program is aimed at delivering counseling support, then a more direct measure of affective outcomes than an achievement assessment is desired.

Valid data are quality data.

3. Investment

A key element frequently cited as basic for achieving quality is the reliance upon and use of the data by the persons responsible for collecting and reporting them. School clerks who never receive feedback or see reports using the discipline data they enter into a computer screen have little investment in the data. School clerks who enter purchasing information into an automated system that tracks accounts and balances have a double investment. They save time when the numbers add up, and they receive praise or complaints if they do not. Whoever is responsible for collecting, entering, or reporting data needs to have a natural accountability relationship with those data. The data providers should experience the consequences of the quality of the data they report.

This may be the most important truism in this paper:

The user of data is the best recorder of data.

4. Certification

Typically, organizations have a set of "official" statistics that are used, regardless of their quality, for determining decisions such as funds allocation or tracking changes over time. These official statistics are needed to provide some base for planning, and the decision makers are challenged to guess how close they are.

Organizations should certify a set of official statistics.

5. Publication

Public reporting or widespread review is a common action cited in the evolution of an information system toward quality.

In every state that has instituted a statewide accountability system, there are stories of the poor quality of the data in the first year. Depending upon the complexity of the system and the sanctions imposed, (either money or reputation) subsequent improvements in data quality were seen.

ESP Insight

Trust must be present for data to be convincing.

The most practical and easily achieved action for impacting data quality is:

Publish the data.

6. Trust

Decision makers refer to the trust and confidence they must have in both the data and the individuals providing the data.

Trust is a crucial component of the working relationship between decision makers and staff within an organization. That trust must be present for data to be convincing. Consultants are used at times to provide that trust and confidence. Decision makers often do not have the time nor the expertise to analyze data. They rely upon someone else's recommendation. Data should be presented by an individual in whom the decision makers have confidence and trust.

Trust the messenger.

These six statements faithfully summarize the insights of professionals who have struggled with data quality within their information systems. They address processes that contribute toward achieving data quality—the dynamics influencing quality within an information system. They do not yet clearly indicate how successful the organization has been in achieving quality. To make that connection, the following hierarchy was developed.

A Hierarchy of Data Quality – Getting to data-driven decision making

This original hierarchy of data quality was designed in the 90's to describe how quality develops and can be achieved. Secretary Rod Paige included the poster Steps for Ensuring Data Quality, Attachment A, which included the hierarchy, in his guidance to states for implementing the data and technology requirements of the No Child Left Behind Act.

ESP Insight

The highest level of quality is achieved when data-based decisions are made with confidence.

The highest level of quality is achieved when data-based decisions are made with confidence. Therefore, several components of quality must be present, i.e., available data, decisions based upon those data, and confidence by the decision maker. Ultimately, quality data serve their intended purpose when the decision maker has the trust to use them with confidence. The traditional virtues of quality (e.g., reliability and validity) form the basis for that trust, but do not ensure it. Accuracy is the traditional characteristic defined within formal information systems architecture. Accuracy begs the question of whether or not the data are worthy of use.

From the observations of organizational quests for quality information systems, the concept of official data has been described. Data are official if they are designated as the data to be used for official purposes, e.g., reporting or calculation of formulas such as for funding schools and programs. At the earliest stages of information systems, the characteristic of being available is the only claim to quality that some data have. The level at the base of the hierarchy is characterized by no data being available.

Examples are provided below to illustrate each level. As you will notice, most of these are from the 80's and 90's when I was managing information systems in a local school district. I was more comfortable using these examples from my own work than more recent ones from our ESP client engagements.

Bad Data

-1.1 Invalid

Bad data can be worse than no data at all. At least with no data, decision makers rely upon other insights or opinions they trust. With bad data, decision makers can be misled. Bad data can be right or wrong, so the actual impact on a decision's outcome may not always be negative. Bad data can result from someone's not understanding why two numbers should not be compared or from errors and inconsistencies throughout the reporting process. The definition of bad data is that they are either:

- Poorly standardized in their definition or collection to the extent that they should be considered unusable, or
- inaccurate, incorrect, unreliable.

ESP Insight

The highest level of quality is achieved when data-based decisions are made with confidence.

An example of bad data occurred when a local high school failed to note that the achievement test booklets being used were in two forms. The instructions were to ensure that each student received the same form of the exam for each subtest. However, the booklets were randomly

distributed each day of the testing, resulting in a mixture of subtest scores that were either accurate (if the student took the form indicated on the answer document) or chance level (if the form and answer document codes were mismatched). This high school was impacted at the time by cross-town bussing that created a very diverse student population of high and low achievers. From our previous analyses, we also knew that an individual student's scores across subtests could validly range plus or minus 45 percentile points. Simple solutions to interpreting the results were not available. (***Empty Bubbles: What Test Form Did They Take?*** D. Doss and G. Ligon, Presented at the American Educational Research Association Annual Meeting, 1985.)

Carolyn Folke, Information Systems Director for the Wisconsin Department of Education, contributed the notion that the hierarchy needed to reflect the negative influence of bad data. In her experience, decision makers who want to use data or want to support a decision they need to make are vulnerable to grasping for any and all available data—without full knowledge of their quality. The message here is look into data quality rather than assume that any available data are better than none.

None

0.0 Unavailable

Before "A Nation at Risk," before automated scheduling and grade reporting systems, and before the availability of high-speed computers, often there were no data at all related to a decision. So, this is really the starting point for the hierarchy.

When a local school district began reporting failure rates for secondary students under the Texas No Pass/No Play Law, one school board member asked for the same data for elementary students. The board member was surprised to hear that, because elementary grade reporting was not automated, there were no data available. (After a long and painful process to collect elementary grade data, the board member was not pleased to learn that very few elementary students ever receive a failing grade and that fewer fail in the lower achieving schools than fail in the higher achieving schools.) (***No Pass - No Play: Impact on Failures, Dropouts, and Course Enrollments,*** G. Ligon, Presented at the American Educational Research Association Annual Meeting, 1988.)

When no data are available, the options are typically obvious—collect some or go ahead and make a decision based upon opinion or previous experience.

However, there is another option used by agencies involved in very large-scale data collections. The Bureau of the Census and the National Center for Education Statistics both employ decision rules to impute data in the absence of reported numbers. Missing cells in tables can be filled with

imputed numbers using trends, averages, or more sophisticated prediction analyses. Decision makers may perform their own informal imputations in the absence of data.

Available

1.1 Inconsistent Forms of Measurement

Poor data come from inconsistencies in the ways in which outcomes or processes are measured. These inconsistencies arise from use of nonparallel forms, lack of standardized procedures, or basic differences in definitions. The result is data that are not comparable.

In 1991, we studied student mobility and discovered that not only did districts across the nation define mobility differently, but they also calculated their rates using different formulas. From 93 responses to our survey, we documented their rates and formulas, and then applied them to the student demographics of Austin. Austin's "mobility" rate ranged from 8% to 45%, our "turbulence" rate ranged from 10% to 117%, and our "stability" rate ranged from 64% to 85%. The nation was not ready to begin comparing published mobility rates across school districts. (***Student Mobility Rates: A Moving Target,*** G. Ligon and V. Paredes, Presented at the American Educational Research Association Annual Meeting, 1992.)

ESP Insight

Equating each state's performance levels with those of NAEP is a popular method for judging the difficulty of assessments across states.

A future example of this level of data quality may come from changes in the legislation specifying the nature of evaluation for Title I Programs. For years, every program reported achievement gains in normal curve equivalent units. Current legislation requires each state to establish an accountability measure and reporting system. Equating each state's performance levels with those of NAEP is a popular method for judging the difficulty of assessments across states.

Full time equivalents and head counts, duplicated and unduplicated counts, average daily attendance and average daily membership are all examples of how state accountability systems must align the way schools maintain their records. Who is not familiar with the "problem" of whether to count parents in a PTA meeting as one attendee each or as two if they have two students in the school?

1.2 Data Collected by Some at Some Times

Incomplete data are difficult to interpret.

In 1994, the *Austin American Statesman* published an article about the use of medications for ADD/ADHD students in the public schools. The headline and point of the story was that usage was much lower than had been previously reported. The person quoted was not a school district employee and the nature of some of the statistics caused further curiosity. So, I called the reporter, who said he had not talked to the District's Health Supervisor and that the facts came from a graduate student's paper. Checking with the Health Supervisor showed that only about half the schools had participated in the survey, some of those with the highest levels of use did not participate, the reporter used the entire District's

membership as the denominator, and the actual usage rate was probably at least twice what had been reported. The reporter's response: "I just reported what she told me."

1.3 Data Combined, Aggregated, Analyzed, Summarized

The highest level of "available data" is achieved when data are summarized in some fashion that creates interesting and useful information. At this point in the hierarchy, the data begin to take on a usefulness that can contribute to a cycle of improved quality. At this point, audiences are able to start the process of asking follow-up questions. The quality of the data becomes an issue when someone begins to use summary statistics.

One of the most dramatic responses to data I recall was when we first calculated and released the numbers and percentages of overage students, those whose age was at least one year over that of their classmates. Schools have always had students' ages in the records. Reality was that no one knew that by the time students reached grade 5 in Austin, one out of three was overage. In at least one elementary school over 60% of the fifth graders were old enough to be in middle school. (The number of elementary retention's began to fall until the rate in the 90's was about one fifth of the rate in the 80's.) (***Do We Fail Those We Fail?***, N. Schuyler and G. Ligon, Presented at the American Educational Research Association Annual Meeting, 1984; ***Promotion or Retention,*** Southwest Educational Research Association Monograph, G. Ligon, Editor, 1991.)

ESP Insight

Data become information when decision-makers can understand them.

When relatively unreliable data are combined, aggregated, analyzed, and summarized, a major transformation can begin. Decision makers can now apply common sense to the information. Data providers now can see consequences from the data they report. This is an important threshold for data quality. In countless conversations with information systems managers and public school evaluators, a consistent theme is that when people start to see their data reported in public and made available for decision making, they begin to focus energies on what those data mean for them and their school/program.

Texas schools began reporting financial data through PEIMS (Public Education Information Management System) in the 1980's. The first data submissions were published as tables, and for the first time it was simple to compare expenditures in specific areas across schools and districts. Immediately, a multi-year process began to bring districts more in line with the State's accounting standards and to ensure better consistency in the matching of expenditures to those categories. When districts reported no expenditures in some required categories and others reported unrealistically high amounts, the lack of data quality was evident. The persistent lack of consistency across districts prompted the Texas Legislature in 2006 to fund a new study and development of a more standardized financial reporting process.

DATA BECOME INFORMATION. *Around this point in the hierarchy, data become information. The individual data elements are inherently less useful to decision makers than are aggregated and summarized statistics. From this point on in the hierarchy, basic data elements are joined by calculated elements that function as indicators of performance.*

Official

2.1 Periodicity Established for Collection and Reporting

Periodicity is the regularly occurring interval for the collection and reporting of data. An established periodicity is essential for longitudinal comparisons. For valid comparisons across schools, districts, and states, the same period of time must be represented in everyone's data.

ESP Insight

Periodicity—an agency must manage the periodicity of its data to understand what is available when.

The National Center for Education Statistics (NCES) has established an annual periodicity set around October 1 as the official date for states to report their student membership. Reality is that each state has its own funding formulas and laws that determine exactly when membership is counted, and most do not conduct another count around October 1 for Federal reporting.

I was called on the carpet by my local superintendent once because a school board member had used different dropout rates than he was using in speeches during a bond election. He explained very directly that "Every organization has a periodicity for their official statistics." That of course is how they avoid simultaneous speeches using different statistics. After working hard with the staff to publish a calendar of our official statistics, I discovered that very few districts at the time had such a schedule. (***Periodicity of Collecting and Reporting AISD's Official Statistics***, G. Ligon et al., Austin ISD Publication Number 92.M02, November, 1992.)

2.2 Official Designation of Data for Decision Making

Finally, official statistics make their way into the hierarchy. The key here is that "official" does not necessarily guarantee quality. Official means that everyone agrees that these are the statistics that they will use. This is a key milestone, because this designation contributes to the priority and attention devoted to these official statistics. This in turn can contribute to on-going or future quality.

Sometimes politimetrics turn out to be better than legacy statistical processes. Every year, our Management Information Department's Office of Student Records issued its student enrollment projection. The preliminary projection was ready in January for review and a final projection for budgeting was ready by March. Here is another example of how the presence of a bond election can influence the behavior of superintendents and school board members. The superintendent gave a speech to the Chamber of Commerce using the preliminary projection. Then our office sent him the final projection. He was not happy with the increase of about 500 in the projection. He believed that created a credibility gap between the figures used in campaigning for the bonds and the budgeting process. So, the preliminary projection, for the first time in history, became the final, "official" projection. The bonds passed, the next year's enrollment was only a few students off of the "official" projection, the School Board was impressed with the accuracy of the projection, and

Austin began a series of four years when all the projection formulas were useless during the oil and real estate bust of the late 80's. The next time the "official" projection was close was when a member of the school board insisted that the district cut 600 students from its projection in order to avoid having to budget resources to serve them.

THE RIGHT DATA MUST BE USED. *At this point, the qualities of accuracy and reliability are required. Moreover, the best data are not quality data if they are not the right data for the job.*

2.3 Accuracy Required for Use in Decision Making

With the official designation of statistics, either by default or intent, their use increases. Now the feedback loop takes over to motivate increased accuracy. The decision makers and the persons held accountable for the numbers now require that the data be accurate.

When we began publishing six-week dropout statistics for our secondary schools, the principals started to pay attention to the numbers. They had requested such frequent status reports so the end-of-the-year numbers would not be a surprise, and so they could react if necessary before the school year was too far along. Quickly, they requested to know the names of the students that we were counting as dropouts, so verification that they had actually dropped out could be made. Having frequent reports tied directly to individual student names improved the quality of the dropout data across the schools.

ESP Insight

The analyses must be appropriate to the question being addressed. The handy process available with an analysis tool may not meet the required assumptions for your data.

THE RIGHT ANALYSES MUST BE RUN. *The quality of data is high at this point, and the decision maker is relying upon analyses conducted using those data. The analyses must be appropriate to the question being addressed.*

*A caution to data providers and audiences: There are times when data quality is questioned, but the confusing nature of the data comes from explainable anomalies rather than errors. We should not be too quick to assume errors when strange results arise. For example, a district's overall average test score can decline even when all subgroup averages rise; students can make real gains on performance measures while falling farther behind grade level; schools can fail to gain on a state's assessment, but be improving. (**Anomalies in Achievement Test Scores: What Goes Up Also Goes Down,*** G. Ligon, Presented at the American Educational Research Association Annual Meeting, 1987.)

Valid

3.1 Accurate Data Consistent with Definitions

Trained researchers are taught early to define operationally all terms as a control in any experiment. Every organization should establish a standard data dictionary for all of its data files. The data dictionary provides a definition, formulas for calculations, code sets, field characteristics, the periodicity for collection and reporting, and other important descriptions. Using a common data dictionary provides the organization the benefits of efficiency by avoiding redundancy in the collection of data elements. Another important benefit is

the ability to share data across departmental data files. (***Periodicity™ User Guide,*** Evaluation Software Publishing, Austin, Texas, 1996.)

The classic example of careless attention to definitions and formulas is *Parade Magazine*'s proclamation that an Orangeburg, South Carolina, high school reduced its dropout rate from 40% to less than 2% annually. Those of us who had been evaluating dropout-prevention programs and calculating dropout rates for a number of years became very suspicious. When newspapers around the nation printed the story that the dropout rate in West Virginia fell 30% in one year after the passage of a law denying driver's licenses to dropouts, we were again skeptical. Both these claims had a basis in real numbers, but each is an example of bad data.

ESP Insight

Authentic assessments have failed the bias test and remain useful for formative but not accountability processes.

The *Parade Magazine* reporter compared a four-year, longitudinal rate to a single-year rate for the Orangeburg high school. The newspaper reporter compared West Virginia's preliminary dropout count to the previous year's final dropout count. (The West Virginia state education agency later reported a change from 17.4% to about 16%.) (***Making Dropout Rates Comparable: An Analysis of Definitions and Formulas,*** G. Ligon, D. Wilkinson, and B. Stewart, Presented at The American Educational Research Association Annual Meeting, 1990.)

3.2 Reliable Data Independent of the Collector

Reliability is achieved if the data would be the same regardless of who collected them.

What better example is available than the bias in teacher evaluations? When Texas implemented a career ladder for teachers, we had to certify those eligible based upon their annual evaluations. The school board determined that they were going to spend only the money provided by the State for career ladder bonuses, so that set the maximum number of teachers who could be placed on the career ladder. Our task was to rank all the eligible teachers and select the "best." Knowing there was likely to be rater bias, we calculated a Z score for each teacher based upon all the ratings given by each evaluator. Then the Z scores were ranked across the entire district. The adjustments based upon rater bias were so large, that near perfect ratings given by a very easy evaluator could be ranked below much lower ratings given by a very tough evaluator. The control was that the teachers' rankings within each rater's group were the same.

Everything was fine until a school board member got a call from his child's teacher. She was her school's teacher-of-the-year candidate but was ranked by her principal in the bottom half of her school, and thus left off the career ladder. The end of the story is that the school board approved enough additional local money to fund career ladder status for every teacher who met the minimum state requirements, and we were scorned for ever having thought we could or should adjust for the bias in the

ratings. (***Adjusting for Rater Bias in Teacher Evaluations: Political and Technical Realities,*** G. Ligon and J. Ellis, Presented at the American Educational Research Association Annual Meeting, 1986.)

3.3 Valid Data Consistent with the Construct Being Measured

The test of validity is often whether a reasonable person accountable for an outcome agrees that the data being collected represent a true measure of that outcome. Validity is the word for which every trained researcher looks. Validity assumes both accuracy and reliability. Critically, valid data are consistent with the construct being described. Another perspective on this is that valid data are those that are actually related to the decision being made.

The local school board in discussing secondary class sizes looked at the ratio of students to teachers in grades 7 through 12 and concluded that they were fairly even. Later they remembered that junior high teachers had been given a second planning period during the day, so their actual class sizes were much higher. Then they moved on to focus on the large discrepancies between class sizes within subject areas to discover that basic required English and mathematics classes can be efficiently scheduled and are large compared to electives and higher level courses. In the end, the school board members became more understanding of which data are valid for use dependent upon the questions they are asking.

ESP Insight

Quality data must be interpretable beyond the local context.

Quality

4.1 Comparable Data: Interpretable Beyond the Local Context

Quality is defined here beyond the psychometric and statistical concepts of reliability and validity. Quality is defined by use. Quality data are those that function to inform decision making. For this function, the first criterion is:

Quality data must be interpretable beyond the local context. There must be a broad base of comparable data that can be used to judge the relative status of local data. We can recognize that there are some decisions that do not necessitate comparisons, but in most instances a larger context is helpful. Each time I read this criterion, I rethink it. However, it is still in the hierarchy because decisions made within the broadest context are the best informed decisions. Knowing what others are doing, how other districts are performing does not have to determine our decisions, but such knowledge ensures that we are aware of other options and other experiences.

ESP Insight

Data-based decisions must

Most states and districts have struggled with defining and reporting their dropout rates. Despite the lofty goal often embraced of having 100% of our students graduate, there is still the need for comparison data to help interpret current levels of attrition. When we compared Austin's dropout rate to published rates across the nation, we found that the various formulas used by others produced a range of rates for Austin from 11% to 32%. Our best comparisons were across time, within Austin, where we had control over the process used to calculate comparable rates. (***Making Dropout Rates Comparable: An Analysis of Definitions and Formulas,*** G.

Ligon, D. Wilkinson, and B. Stewart, Presented at The American Educational Research Association Annual Meeting, 1990.)

4.2 Data-Based Decisions Made with Confidence

The second criterion is:

> Data-based decisions must be made with confidence, at least confidence in the data. This is the ultimate criterion upon which to judge the quality of data--do the decision makers who rely upon the data have confidence in them. Assuming all the lower levels of quality criteria have been met, then the final one that makes sense is that the data are actually used with confidence.

This is a good time to remind us all that confidence alone is not sufficient. One reason the construct of a hierarchy is useful is that each subsequent level depends upon earlier levels.

> A local district's discipline reporting system had been used for years to provide indicators of the number of students and the types of incidents in which they were involved. The reports were so clear and consistent that confidence was high. As part of a program evaluation, an evaluator went to a campus to get more details and discovered that only about 60% of all discipline incidents were routinely entered into the computer file. The others were dealt with quickly or came at a busy time. No one had ever audited a school's discipline data. On the other hand, the dropout and college-bound entries into a similar file were found to be very accurate and up-to-date.

Steps for Ensuring Data Quality

The poster in Attachment A details the six steps for ensuring data quality. Each is stated as a question to form a checklist.

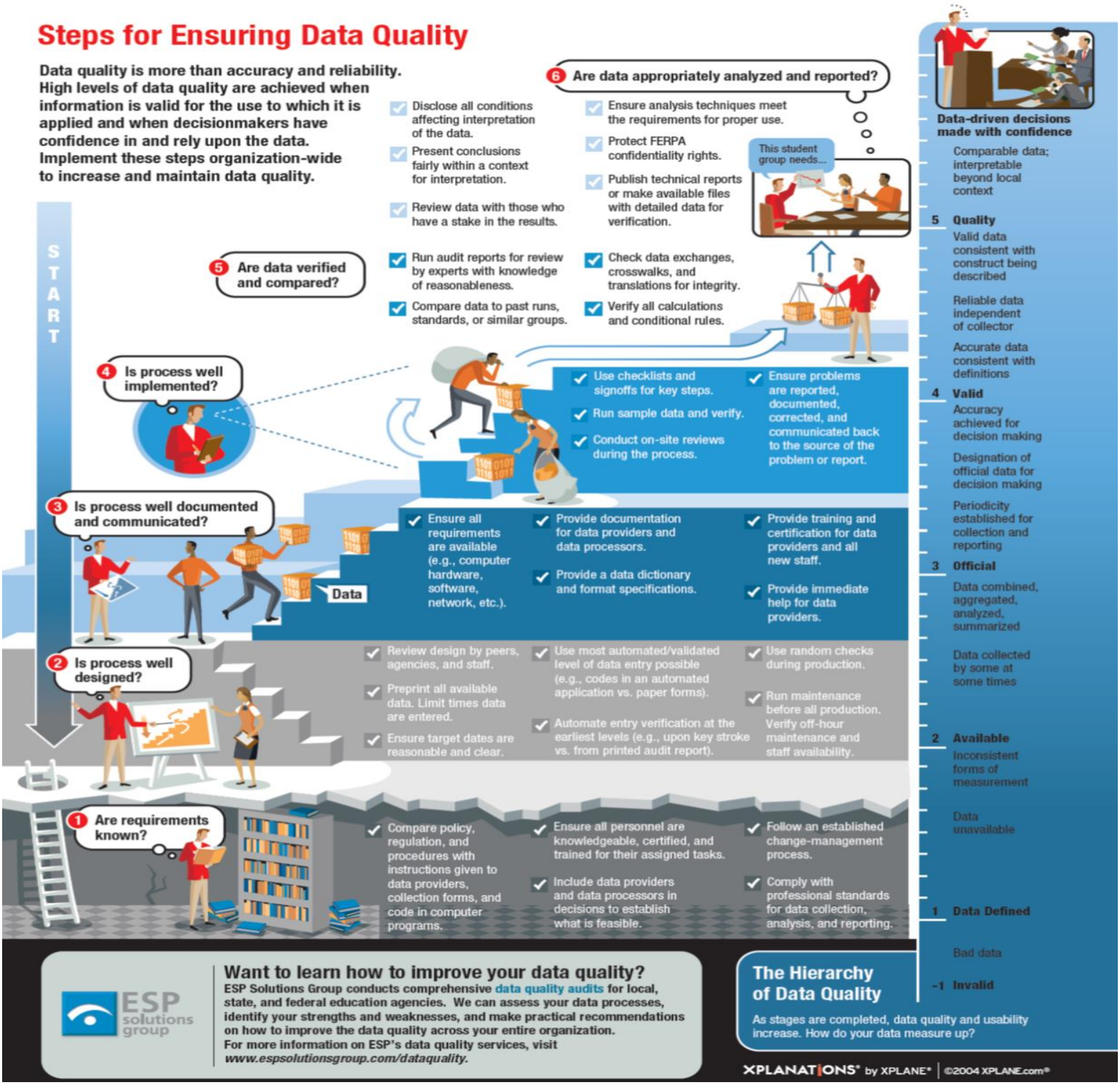
1. Are requirements known?
2. Is process well designed?
3. Is process well documented and communicated?
4. Is process well implemented?
5. Are data verified and compared?
6. Are data appropriately analyzed and reported?

To supplement these steps, ESP compiled lessons learned and advice into the Data Quality Boot Camp provided in Attachment B.

Conclusion

The hierarchy and the steps for ensuring data quality were a convenient way to think through what makes for quality data. Reality is that our information systems will not fall neatly into one of the levels of the hierarchy. In fact they may not often evolve sequentially through each level. At any point in time, their levels may shift up or down. What is useful here is that the hierarchy describes the characteristics of relatively low and relatively high levels of data quality.

Attachment A



Attachment B

Data Quality Boot Camp – Understanding the principles of data quality

Ready to go through a boot camp for data quality? The basics of ensuring and maintaining quality data throughout an information system have been gleaned from our ESP experts and summarized below.

Data quality, the basics:

1. Get data right from the start.
2. Keep them right at every step.
3. Give people help to do this.

The next person in line can't fix the last person's errors as easily as that person can.

Poor data quality, the culprits:

1. Missing data
2. Incorrect data
3. Late data

Most vulnerable times for data:

1. Entry
2. Exchange

The Four Great Truths about Data Quality:
Data quality is highest when…

1. The data providers know what's expected.
2. The data providers use the data themselves for their own work.
3. Everyone, everywhere checks the data.
4. The data are available and used.

ESP Insight

The Data Quality Boot Camp is not merely platitudes. These are insights from across every state and a full range of information system architectures.

Principles of Data Quality

Data quality abides by some well-tested principles. The fact that these are not widely known is a shame.

The *Expectation* Principle of Data Quality

- Data quality can only be achieved when the expectations are clear.

  Documentation of data definitions, codes, and business rules is essential. Metadata—be sure the data providers have been told.

The *Use* Principle of Data Quality

- Data quality matters when the data are used by the person collecting and reporting the data.

The high school registrar is the law when it comes to official transcript data. The registrar must certify that the records are complete, accurate, and official, so nothing gets out without scrutiny.

- Data quality requires all data handlers to check their own data.
  No one can spot errors and omissions in your data better than you. Don't pass along your errors and expect the next person to find and correct them.

The *Comparability* Principle of Data Quality

- Data quality matters when the data are compared.
  Is your school's attendance rate really lower than your rival's? Are you treating excused absences the same way?

The *Hierarchical Norm* Principle of Data Quality

- Each institution is a subject of a higher institution and an authority for a lower institution.
- Every data element an authority chooses to define must be defined the same by all lower institutions.

Institutional Hierarchy
US Department of Education
State Education Agency
Local Education Agency (District)
School
Employee

Notice that the individual tasked with providing the data is not an authority for the data.

The *Transformation* Principle of Data Quality

- A subject institution may define a data element differently from its authority only to the extent that the data element can be derived from or transformed into the precise definition of the higher authority.
  Keep more detail, use your own codes, but be sure you can transform it all to the required categories.

The *Transformation Burden* Principle of Data Quality

- Part A: The burden to transform is solely the burden of the subject institution.
- Part B: This burden compels the subject institution to comply with the standard of the authority.
  It's just easier to do it right the first time. Why have to transform your codes if you can use the standard ones from the beginning?

ESP Insight

These Principles of Data Quality get translated into specific steps in Part II of this series on data quality.

ESP Insight

Creativity, forgiveness, procrastination, and delegating upward are not principles in a quality data process.

The *Monkey on My Back* version of the Transformation Burden Principle of Data Quality

- Data Provider: I can get them to clean the data because they are the ones who want it anyway.
- Data Requestor: I'm the one who needs these data, so I have to clean them up if they won't.

    This is the root cause of so much pain. The requestor is the enabler. If rules are enforced from the beginning, data providers get the message that they can do it right now or do it again before the requestor will take it.

The *Invented Here* Principle of Data Quality

- Competes with the Transformation Burden Principle.
- As the local expert, I know how we should define our data.

    Not a team player, this know-it-all. The rules must be enforced even with the legendary staff members who have been around since the beginning of computer time.

The *Vendor Rules* Principle of Data Quality

- When we chose our vendor, we chose our data standards.

    No, no, no. Vendors want your business and your reference. Leverage that to get what you need.

The *Inertia* Principle of Data Quality

- If we change to use the authority's standard, we have to retrain everyone and reconfigure all our software.

    Yes, you do. Do it.

What does this mean for me?

- If you follow the authority's rules, burden is lower.
- If you change the rules, you have to re-work your data for reporting.

What does this mean for data quality?

- If people follow the rules, quality is higher.
- If people change the rules, quality is not achieved.

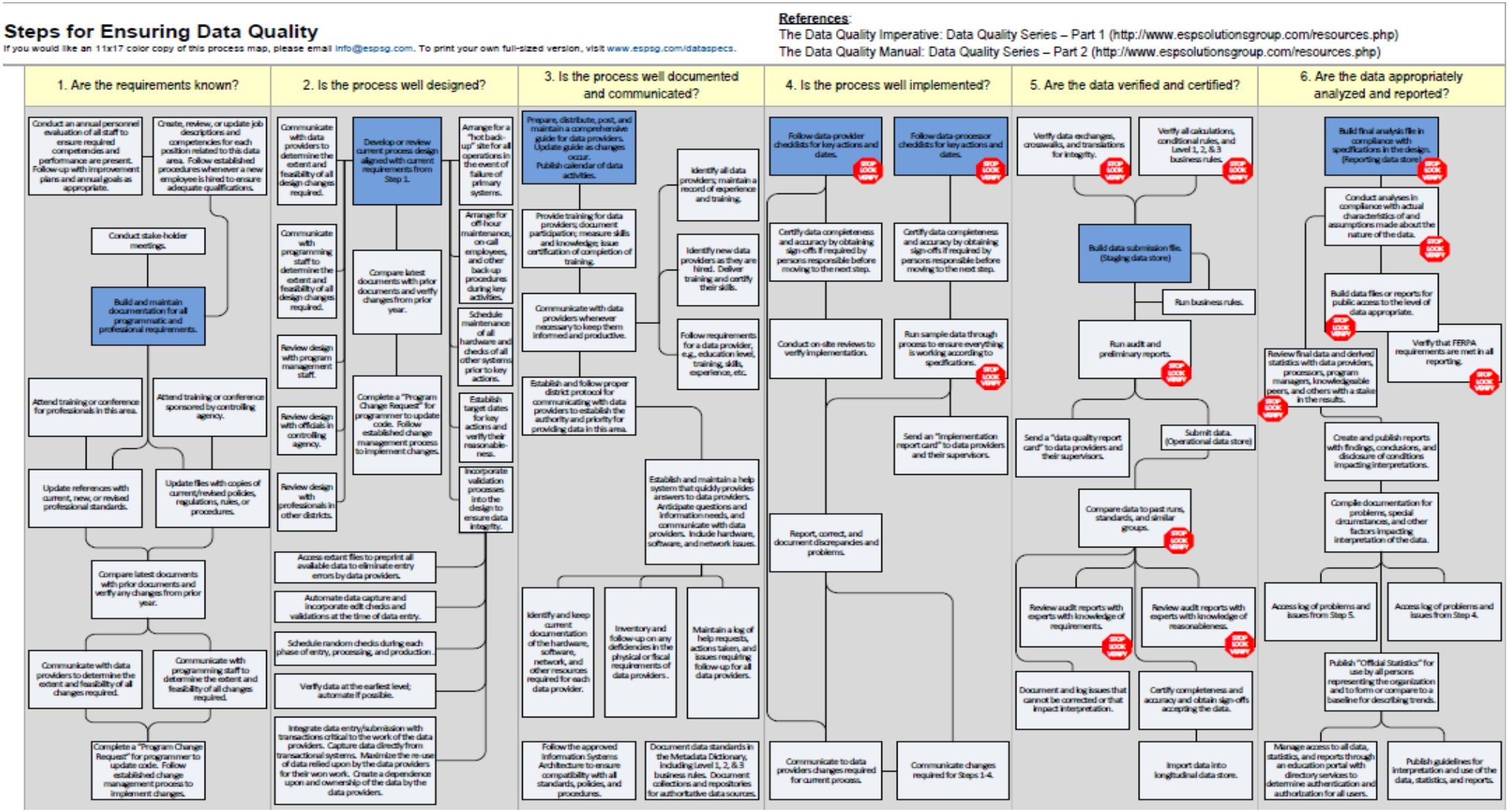
The unfortunate truth about reporting quality data:

- If you do something well the first time, people will not appreciate how difficult it is to do.

The redeeming factor:

- Getting data right from the start is difficult. However, providing clean, timely data is greatly appreciated by the collector.

Attachment B: Process Illustration of Quality



About ESP Solutions Group

ESP Solutions Group provides its clients with *Extraordinary Insight™* into P20W education data systems and psychometrics. Our team is comprised of industry experts who pioneered the concept of "data-driven decision making" and now help optimize the management of our clients' state and local education agencies' information systems.

ESP personnel have advised school districts, all state education agencies, and the U.S. Department of Education on the practice of P20W data management. We are regarded as leading experts in understanding the data and technology implications of NCLB, SIF, ED*Facts*, CEDS, state reporting, metadata standards, data governance, data visualizations, and emerging issues.

Dozens of education agencies have hired ESP to design and build their longitudinal data systems, state and federal reporting systems, metadata dictionaries, evaluation/assessment programs, and data management/analysis and visualization systems.

To learn how ESP can give your agency *Extraordinary Insight* into your P20W education data, contact us at (512) 879-5300 or info@espsg.com.

This document is part of *The Optimal Reference Guide* Series, designed to help decision makers analyze, manage, and share data in the 21st Century.

ESP Solutions Group

(512) 879-5300
www.espsolutionsgroup.com