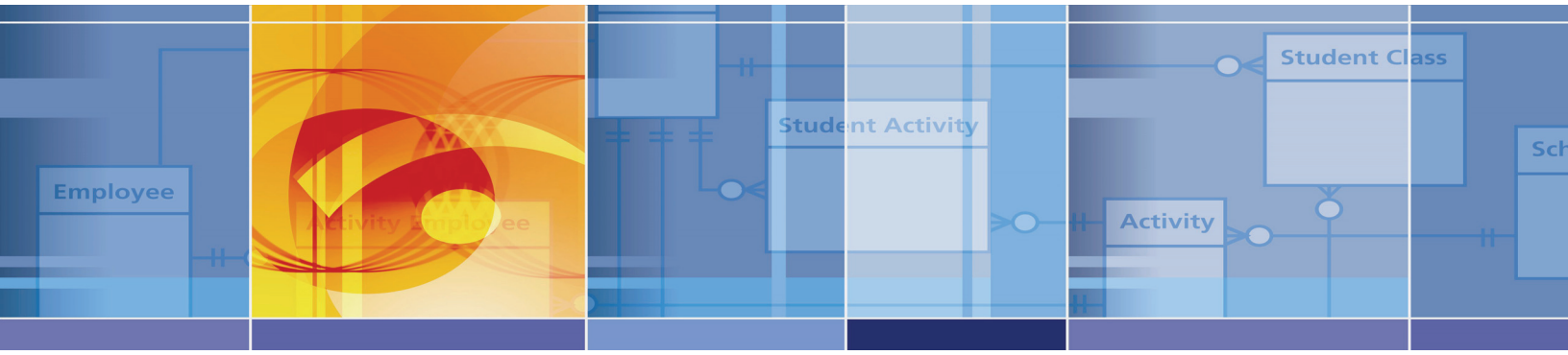


The Optimal Reference Book:

Aligning Indicators and Actions

Extraordinary insight™ into today's education information topics

ESP Solutions Group

Table of Contents

About ESP Solutions Group

ESP Solutions Group provides its clients with *Extraordinary Insight*™ into PK-12 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.

ESP personnel have advised school districts, all 52 state education agencies, and the U.S. Department of Education on the practice of K-12 school data management. We are regarded as leading experts in understanding the data and technology implications of the **No Child Left Behind Act (NCLB)**, **ED*Facts***, and the **Schools Interoperability Framework (SIF).**

Dozens of education agencies have hired ESP to design and build their student record collection systems, federal reporting systems, student identifier systems, data dictionaries, evaluation/assessment programs, and data management/analysis systems.

To learn how ESP can give your agency *Extraordinary Insight* into your PK-12 education data, email info@espsg.com.

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

About the Author

Glynn D. Ligon, Ph.D., President and CEO

Dr. Ligon, the president and chief executive officer of ESP Solutions Group, is a nationally recognized expert on issues relating to student record collection and exchange, data quality, data reporting, and large-scale system design.

The National Center for Education Statistics, the U. S. Department of Education and over 25 state education agencies have consulted with Dr. Ligon on various areas of his expertise. He has a Ph.D. in Educational Psychology, Quantitative Methods from The University of Texas at Austin and is licensed to teach in the State of Texas.

Prior to starting ESP in 1993, Dr. Ligon directed the Austin (TX) Independent School District's information and technology organization. As the executive director of management information, he led the district's efforts in developing and reporting on district-wide program evaluations, many of which won national awards from the American Educational Research Association. Dr. Ligon was also a leader in the advent of SPEEDE/ExPRESS, the EDI standard for the exchange of electronic student transcripts.

From 1992 to 2000, he served as a member of the U.S. Department of Education's Planning and Evaluation Services Review Panel. Dr. Ligon's whitepapers; *A Technology Framework for NCLB Success* and *Steps for Ensuring Data Quality* are prominently featured within the U.S. Department of Education's 2005 National Education Technology Plan, meant to help motivate and incite technology-driven transformation in education.

At the beginning of his career, Dr. Ligon taught in predominantly Spanish-speaking schools near the Texas-Mexico border. He is an experienced evaluator of Title I, Migrant, compensatory education, and bilingual education programs.

Actions Speak Louder than Data

Foreword

By Barbara S. Clements, Ph.D., ESP's Vice President for Education Services

This Optimal Reference Guide represents a return for ESP to its roots. The first ESP product, SuccessFinder™, was a software application that conducted sophisticated analyses of student performance data. Apparently, in 1991 it was "ahead its time" because a big seller it was not. Everyone who heard about SuccessFinder said that it must be an excellent product because Glynn Ligon developed it, and he knows so much about data analysis and reporting. But eventually we shelved the product, though not what we learned in building it.

An important thing we learned in trying to market SuccessFinder was that most potential users (i.e., school districts or state education agencies) did not have the data they needed to use the application or they did not trust the data they had. It became clear to Glynn and his ESP colleagues that the data infrastructure in school districts and state education agencies needed substantial work before the data could be appropriately used.

In a decision support system project we did for the U.S. Department of Education's Office of the Chief Information Officer, we had the opportunity to look at what states and school districts were doing with their data, and how the data were being used. Despite seeing how much data were being collected, we rarely saw any "actionable" reports being produced. One school district showed us how they had learned that attendance was down on the day before a holiday and on Fridays in general. But they couldn't tell us how they were using those data. And so it went.

Well, now a substantial amount of work has been done on the data infrastructure in school districts and state education agencies around the nation. More data are being collected primarily in individual student records, which offer more flexibility for analysis. In addition, much work has been done on the quality of the data being collected. So maybe it's time to drag out what we have learned about developing reports from the data and help educators make more appropriate decisions about students and schools.

In this paper, Glynn has not only dusted off and presented what we learned with our SuccessFinder experience, but also provided a more insightful and extensive understanding of how education data can be used effectively by educators. As usual, he is ahead of his time.

Introduction: Time to Change our Mind(set)

When I was a school district employee responding to *ad hoc* requests for information from principals, superintendents, and parents, I was struck by how many times the requestors didn't get back what they really wanted. After tiring of re-running analyses, I realized that the requestor's mindset was turned around. Instead of walking in and asking for specific data, they should have been telling me what they wanted to do with the data. That "aha" made life so much easier.

This Optimal Reference Guide turns around how educators are using their decision support systems. Educators look at the available data and say, "What can I learn from these data?" Instead they should be asking first, "What do I need to do that can be informed by data?"

We are doing a poor job of informing decisions with data. Certainly there are many decision support systems, reporting applications, and query tools available. However, the reports I've seen are disappointing. That's not a significant conclusion from someone who has seen demonstrations of 10 major reporting applications over the past six months.

ESP Insight
Imagine reports that hit your monitor at the time an action can be taken—an action informed by the data.

How many ways can we disaggregate, drill down, and graph the same data? In the 21st century, educators are still limited to readily available data. How are we going to define the data we should be collecting? In the end, the majority of today's reports are still merely counts. I'm looking for more than that. Imagine reports that hit your monitor at the time an action can be taken—an action informed by the data.

This paper pushes us to think beyond the limits of our current data, beyond the edge of descriptive reports that display counts, all the way to relationships within the data that reveal what needs to be done. This is not simple. Groups I have worked with progress quickly from asking for tables of counts to asking for lists of students that share a characteristic that indicates action. That's leaping in the right direction. The gap between knowing there are 45 overage students who failed the mathematics assessment to getting a list of them by classroom is large…but not nearly huge enough. We need to give teachers that list with a plan of action.

Imagine getting lists of students who are likely to drop out this week because their pending report card will tell them their grades are so low that they must make all A's for the first time in their lives next reporting cycle or they will fail too many classes to graduate. Why should these students stay through Monday? Why should they come back next fall?

ESP Insight
*An **action report** is one that informs the user of an opportunity or a requirement and suggests what action should be taken.*

Imagine getting an alert through the education portal on your monitor that the new student who is coming down the hallway to enter your classroom has a hearing loss in the right ear and needs to sit on your left side in the classroom, or has to take a retest of the state assessment in reading next week because three specific standards were not met on the last assessment, or excels in group work rather than individual tasks.

This paper defines a framework for designing and producing action reports. An action report is one that informs the user of an opportunity or a requirement and

suggests what action should be taken. This action report framework does not ignore the reality that many of our reporting requirements are to comply with legal or funding mandates or to simply publish information in a profile.

Types of Reports

The action report framework developed is based upon a taxonomy of five report types. Don't be confused by the use of the word action associated with some reporting that is not traditionally considered very action oriented. I will propose that even filling out a required compliance report is an action. Even creating an annual statistical report or profile is an action. These are not as exciting as reports that reveal which reading program produces greater gains, but they are significant duties in the education world and must be fulfilled responsibly.

The report types are:

1. **Decision**: A judgment is required about the appropriate action to take.

   This is what data driven decision making (D3M) is all about. There may be no law, policy, directive, or requirement that demands the decision maker use real data, but a professional knows that the best decisions are informed ones.

2. **Compliance**: A report is required to meet a mandate.

   These are the reports that try our patience. The requirements are established by someone else, and the burden to create the report is high. They have to be done to get funding, be accredited, or follow a law.

3. **Profile**: A descriptive report is provided (school report card, web profile).

   At times, people just want information. School profiles have become common, basic statistics appear on web sites, and general information that is frequently requested is made handy.

4. **Analysis or *Ad Hoc* Query**: An unanticipated request or a research/evaluation question is answered.

   The nature of *ad hoc* analyses is that we do not anticipate the need for them—or we do not have the resources or time to prepare for them in advance. Research and evaluation questions may be one-time events or too sophisticated to put into a scheduled production mode.

5. **Audit**: A statistical report or list is needed to monitor the functioning of a system or process.

   At times we just need to know if something is working well. Financial processes are not the only ones that require us to monitor transactions frequently.

ESP Insight
The same data may inform all five types of reports.

How do these types overlap? We should not get too focused on the taxonomy. In fact, the action framework developed and described here relies more on your requirements for an individual report than it does on the type of report. The same data may inform all five types of reports. The formatting of the report, the access media, and the timing of each report is influenced by the type.

A report can mix purposes and types. However, especially with decision reports, delivering a simple message is helpful to the user. Producing more, single-purpose reports rather than combining purposes for a report with extensive information is recommended. This simplicity also provides a report design that better targets giving a specific audience only the information appropriate. This targeting results in reports that are easier to interpret and use.

Data are NOT Actionable

We can all agree that data are not actionable. Despite the common expression "actionable data," data are too low level, too simple. We need action reports. However, even more specific, we need indicators on those reports that represent the information that really informs our decision. Then beyond the indicator, we need to establish thresholds on the indicator that determine the action required at each level measured by the indicator.

ESP Insight
Despite the common expression "actionable data," data are too low level, too simple.

This paradigm led to the development of a logical process, which is called the **D3M Action Report Framework™**. There are 10 steps that lead us from the statement of an action to be taken through the production of a report that informs that action.

1. Describe a decision, compliance requirement, profile need, analysis desire, or audit need.

   One of the most essential insights we were taught in our graduate inferential statistics courses was that every experiment must begin with a clear statement of the question to be answered. This is even more essential in the real world because we are too often accepting of available data or statistics rather than independently determining the data we actually need.

2. Define an indicator.

   Indicators give us quick guidance for forming opinions. Indicators are data points that inform our judgment about the status of an entity (e.g., individual, program, or organization). We rely upon indicators to give us facts—even if those facts are representations of opinions from surveys.

3. Set a threshold.

   The threshold sets the level on the indicator at which action is needed. There may be multiple thresholds, all indicative of different actions.

4. Identify the required data elements.

   The indicator and the threshold determine what data are required. The data must be operationally defined in the organization's data dictionary.

5. Identify the data collection that gathers the required data.
   Within an organization, the identified data must be collected to match both the operational definition and the periodicity required to determine levels on the indicator.

6. Identify the data repository that is the authoritative data source for the required data.

   There must be an available, trustworthy, and authoritative source for the required data. To authenticate the quality of the data, the provenance of the data must be established. For those who do not watch the phenomenon called "The Antiques Roadshow," but view instead any one of a dozen legal dramas, the data equivalent of provenance is chain of possession. Who or what system has handled the data from its initial entry through all the exchanges to a final authoritative data store from which an action report will be produced?

7. Describe the Action Report™ to be generated to inform the action.

   To plan for the production of the actual report, several characteristics must be known. These include the report's title, periodicity (schedule), office responsible, media, confidentiality level, etc.

8. Determine the action to be taken at each threshold on the indicator.

   This is the step where the process goes beyond the typical report. This is where an action report is differentiated from a compliance report or profile. Each threshold level that has been identified needs to be aligned with the action that needs to be taken.

9. Design the output format for the report.

   I often caution educators to begin with the report they eventually want to see. However, the prior eight steps are necessary to get to that report ideal. Now is the time to sketch out the actual report to be generated—in the format that supports its use.

10. Access the Action Report™.

    This is the action step. The intended user of the report must get it. This can be by "push" (someone sends the report to the user) or by "pull" (the user requests the report). The point of many action reports is to alert the user that something needs to be done, not to expect the user to be aware of action needed. Therefore, pushing an action report to the user should always be the first consideration.

ESP Insight
Download the Max Yield Data presentation at www.espsg.com.

This process assists an organization in collecting data that are valuable—and saving the time and effort to collect less valuable data. Max Yield Data™ refers to data that everyone agrees are worth the effort to collect, store, and report. (Download the *Max Yield Data* presentation, available at http://www.espsg.com/resources.php.) However, even Max Yield Data must be vetted against their value for informing a decision, determining an action, or answering a question.

One criterion for Max Yield Data is that they are required to inform a decision, specifically a decision that results in action being taken. We already debunked

actionable data as a misnomer because data, in their raw form, are seldom interpretable in a decision-making situation. Statistics, metrics, analytics, and indicators are the ingredients of which decisions and actions are made. These derived values are only useful when they are presented in a report that is formatted and available at the moment it is needed and the user is ready to make a decision.

If you are like me, you've seen enough reports. Tables of numbers, often beautifully presented in graphical forms—in living color. However we can be left thinking—so much data, but so little that I would actually use to form a decision. There are several reasons that have arrived at this point. The most significant one is that we have rushed to put all the data we can find into our data warehouses without a thorough vetting of which data rate the designation of Max Yield Data.

The Action Report Taxonomy

The five types of action reports are detailed in the taxonomy chart below.

Type	Decision	Profile	Compliance	Analysis (*Ad Hoc*)	Audit
Purpose	To trigger an action, inform a decision, or answer a question	To provide general descriptions	To meet a mandate	To answer *ad hoc* questions or inform research and evaluation	To monitor processes
Content/ Scope	Targeted for a specific response	Matched to the audience	Specified by mandating agency	Determined by research question	Metrics describing process
Audience	Individuals who are responsible for taking the action	General audiences that have a broad range of interests	Governance body	Requestor or analyst	System administrator
Media	Quick delivery media, e.g., e-mail, web portal	Stable, official media, e.g., web page, printed report	Mandated media, e.g., data file	Determined by audience	Electronic
Periodicity	Determined by when the action will occur	Determined by availability of the data	Determined by mandating agency	On demand	Continuous, on demand
Analytics	Derivation of an indicator	Statistics and text as desired	Statistics or unit records as mandated	Parametric and nonparametric inferential and descriptive statistics as appropriate	Metrics descriptive of the processes
Data Quality	Ranges from high stakes, high quality demand to soft heads up for possible issues	Court of Public Opinion	May be audited, compared to other reporting agencies	Dependent upon the practices of the analyst	Transactional data quality is imbedded in the analytics
Example Reports	Alert of Students At Risk for Dropping Out	No Child Left Behind Annual Report Card	USED Consolidated Report for Title I and Other Programs; Office for Civil Rights Report	Annual Evaluation of Alternative Reading Programs	Daily Meals Served Report for the National School Lunch Program

Attachment A is an example of an Action Report™ for a school or local education agency (LEA). There are certainly more complex indicators, thresholds, and reports than one focusing on attendance, but this example is intended merely as an illustration. For a state education agency, a more typical example may be a report that identifies students who have been reported as dropouts by one district but who show up on an enrollment report for another. Reclassifying these non-dropouts as

transfer students improves the accuracy and potentially the annual yearly progress (AYP) status of the first school.

The example in Attachment A could have been filled in using the D3M Action Report Framework™ as a template or the metadata descriptions could be entered into DataSpecs™ and the Action Report™ description generated from that database. (DataSpecs™ is a data dictionary tool developed by ESP Solutions Group to document an organization's data standards, collections, repositories, and reports.)

Interpreting and Using Action Reports

Action reports should come with consumer warnings. The reliability of the indicators, statistics, and counts should be explicit.

Type 1 errors are more desirable than type 2 errors. We create less harm over-identifying students than we do when we miss one who really needs attention.

The value added in terms of statistical significance and educational importance of the actual differences teased out by popular sophisticated techniques such as hierarchical linear models (HLM) must be clearly presented. I have seen instances where these impressive analyses are run only to provide miniscule information beyond what we already knew from more straightforward and, quite frankly, more understandable analyses.

ESP Insight
HLM may only tease out miniscule information beyond what we already know.

Growth measures are becoming not only popular but politically mandatory. I would also caution that growth, as compelling as the rationale is to consider it, typically does not reverse our prior judgment of success for individual schools. In other words, a low-performing school usually is one that is delivering low growth as well. However, in this instance, we should be alert for the exceptions. Those exceptional schools that can produce academic growth within a low-performing student population are worthy of identification.

So here are some caveats I have found to be useful when interpreting reports.

- Don't discount the value of an education professional's opinion of a student's status.
- Don't make the mistake of placing too much importance on variables that make too little difference. Occam's Razor tells us that all things being equal, the simplest solution tends to be the best one.
- On the other hand, accept that the interaction of all variables might make the outcome being predicted almost impossible to know.
- Don't place too much trust in averages. A subgroup's average assessment score may not have been made by any single student in the school.
- Look for the quiet, forgotten student who is not being engaged or involved by any teacher, staff member, or other students. That student may be at risk regardless of never being on an action report.
- Perpetual bubble students who do not fall into any one risk category may on balance be at risk.
- If a variable is not measured very precisely, then it likely is adding noise not clarity in a data driven decision making process.
- Given these incomprehensible limitations, don't give up on data.

ESP Insight
Don't give up on data.

Over the years, I have read and participated in setting numerous strategic goals—aligned with an organization's mission. If I were to nominate the most representative goals from across all these efforts, they would be translated into these indicators of success for students.

- What is the quintessential indicator for education organizations?
  - Graduation from high school
- What is the penultimate indicator?
  - Enrollment in higher education
- What is the societal indicator?
  - Graduates who are productive citizens

The importance of indicators in the action report arena cannot be overstated. Analyzing indicators, however, is beyond the scope of this paper. A second Optimal Reference Guide will follow with a thorough analysis of education indicators.

Actions

In all the discussions of and calls I've heard for actionable data, I've never seen a list of actions someone needs to take when the data are available. So, on a recent flight, I began making a list of over 100. When I scanned this list, the first thought I had was, "Some of these questions demand different data than we find in the standard data warehouse. Some of those data are not even collected by most information systems."

ESP Insight
Action reports definitely encompass the administrative services of a school system as well as the instructional services.

To ensure that we do not restrict our thinking of action reports to assessment and enrollment data, I'll point out that our school systems are some of the largest employers, food service providers, transportation systems, facilities managers, financial institutions in the nation. Action reports definitely encompass the administrative services of a school system as well as the instructional services.

There are a surprising number of verbs that describe the action educators take. A few are…

1. Deciding
2. Recommending
3. Approving
4. Choosing
5. Selecting
6. Determining
7. Assigning
8. Aligning
9. Identifying
10. Alerting
11. Fixing
12. Voting
13. Forming (an opinion)
14. Counseling
15. Disciplining
16. Honoring
17. Scheduling
18. Presenting

19. Writing (a report)
20. Grading (an assignment, determining promotion, determining graduation eligibility)
21. Reporting (compliance)
22. Describing (profiling)
23. Researching
24. Evaluating

The sequence of actions for an educational intervention is:

Design

1. Identifying a decision or an action that is driven by data.
2. Defining the indicator used to trigger or inform the action.
3. Establishing a threshold—at what value on a scale is action required?
4. Identifying the data elements required to derive the metric/indicator.
5. Identifying the collection for acquiring the data.
6. Identifying the authoritative data repository where the data can be accessed.
7. Designing a report or alert to go to the person(s) who will take action.
8. Determining an action that will be taken at each threshold level on the indicator.
9. Designing the output format.

Action

10. Collecting, analyzing the data.
11. Identifying individuals or groups passing (or not passing) that threshold.
12. Alerting the right people to respond by providing them the report.
13. Determining a plan of intervention/action.
14. Taking that action.

Evaluation

15. Evaluating the outcome or effectiveness of the action taken.
16. Recommending changes or continuation of the intervention.
17. Modifying 1-9.
18. Repeating 10-17.

D3M is a complex process.

Student Performance Reports

Student performance reports get the most attention in education these days. The standard reports provided by a state's assessment vendor are descriptive--lists of students by subgroup or classroom, percents of students by proficiency levels, and maybe even some old-fashioned average scale scores. Many education agencies have purchased or developed reporting software packages that create OLAP cubes or flat analysis tables from which standard reports or *ad hoc* queries can be run. The training and understanding of the data that are required for a typical educator to use these systems is often too great—not unrealistic, merely requiring time that just isn't available.

ESP Insight
The training and understanding of the data that are required for a typical educator to run ad hoc queries is often too great.

Adding to this conundrum is the reality that analyzing and reporting assessment results requires a thorough understanding of both psychometrics and the actual

assessment. A busy educator may not have the time to learn what the assumptions are underlying the data, what changes have occurred from one year to the next in the scaling, the inclusion/exclusion rules that impact the availability of scores, or the proper way to account for missing data. That's all before the educator gets to the point of measuring the reliability of any statistics generated in an *ad hoc* report.

Over the years, my ESP colleagues and I have created a chart of the questions that people ask about performance results. Each question requires its own type of analysis and interpretation. Now with the D3M Action Report Framework™, we can align each with an example of the action that would be associated with each question.

Attachment B details examples of decisions that are based upon data from student performance measures. The charts describe each question in these terms:

- Decision to be informed
- Question to be answered
- Level of analysis (e.g., individual student or group)
- Narrative of the question for the level of analysis
- Descriptive statistics required
- Actual difference determined
- Statistical significance measure
- Effect size measure
- Analysis of why differences were found

ESP Insight
The major shortcoming of most decision support systems and their reporting tools is that they provide mostly descriptive statistics.

Two perspectives are presented in each chart. The first takes the perspective of informing decisions using status, meaning a single point-in-time measurement. The second perspective uses trend data, meaning how performance changed over time. Growth measures would be in this second category.

These charts illustrate how many questions are not answered by simple tables of data. In fact, the major shortcoming of most decision support systems and their reporting tools is that they provide mostly descriptive statistics.

Conclusion

The D3M Action Report Framework™ was developed to provide educators with a comprehensive process for getting reports they really can use. The process turns out to be rather complex with many moving parts. Not only do compelling reports need to be designed, they also need to be matched to the required data to fill them.

I was handed one of the first Migrant Student Record Transfer System (MSRTS) reports in 1971 while I was teaching at Mission Migrant School in Mission, Texas. Those green-bar, ledger-size printouts were something to behold at the time. Unfortunately, they arrived the last week of school with the requirement that I fill in the data for the school year just ending. Then the sheets went to Little Rock, where they were keyed into the MSRTS system for delivery at the end of the next year to someone else. Useless? Yes.

Ironically, while many of the migrant families have become more stable over the years, the general population has become more mobile.

Mobility, accountability, and professional responsibility are merely three of the compelling reasons to drive our decisions about students with data. I am still working to improve our education reports. I believe we now know what to do.

Attachment A: D3M Action Reports™

(a component of DataSpecs™)—An LEA/School Example

D3M Action Reporting Framework™

This planning framework provides the structure for specifying how actionable data will be provided in a D3M Action Report™ to inform a specific decision, compliance report, profile, analysis, or audit report.

This Framework may be completed from this template or generated as a standard metadata report from DataSpecs™.

NAME OF ACTION:

Attendance Alert for Students At Risk of Truancy

Step 1a. Describe an ACTION.

Describe a decision or question that must be informed by data. *(DataSpecs contains a library of decisions, compliance reports, profiles, analyses, or audit reports.*

Action	Action Identifier in DataSpecs	Description
Intervene for a student with a history of truancy	A0001	A teacher or counselor should intervene early, contact parents, whenever a student with a history of excessive absences first displays the same pattern in a new school year.

Step 1b. Categorize the report by DECISION (D), PROFILE (P), COMPLIANCE (C), ANALYSIS (An), or AUDIT (Au).

The report type helps determine how the Action Report is designed and managed.

In each step, whether the step is required (R) or optional (O) is shown for each report type, e.g., **D = R P = O C = O An = O Au = O**.

Type	Decision	Profile	Compliance	Analysis (*Ad Hoc*)	Audit
Purpose	To trigger an action, inform a decision, or answer a question	To provide general descriptions	To meet a mandate	To answer *ad hoc* questions or inform research and evaluation	To monitor processes
Content/ Scope	Targeted for a specific response	Matched to the audience	Specified by mandating agency	Determined by research question	Metrics describing process
Audience	Individuals who are responsible for taking the action	General audiences that have a broad range of interests	Governance body	Requestor or analyst	System administrator
Media	Quick delivery media, e.g., e-mail, web portal	Stable, official media, e.g., web page, printed report	Mandated media, e.g., data file	Determined by audience	Electronic
Periodicity	Determined by when the action will occur	Determined by availability of the data	Determined by mandating agency	On demand	Continuous, on demand
Analytics	Derivation of an indicator	Statistics and text as desired	Statistics or unit records as mandated	Parametric and nonparametric inferential and descriptive statistics as appropriate	Metrics descriptive of the processes
Data Quality	Ranges from high stakes, high quality demand to soft heads up for possible issues	Court of Public Opinion	May be audited, compared to other reporting agencies	Dependent upon the practices of the analyst	Transactional data quality is imbedded in the analytics
Example Reports	Alert of Students At Risk for Dropping Out	No Child Left Behind Annual Report Card	USED Consolidated Report for Title I and Other Programs; Office for Civil Rights Report	Annual Evaluation of Alternative Reading Programs	Daily Meals Served Report for the National School Lunch Program

Step 2. Define an INDICTOR.

D = R P = O C = O An = O Au = O

Describe the indicator that determines the appropriate action, decision, or answer. *(DataSpecs™ contains a library of predefined indicators.)* The level of analysis, e.g., student, class, teacher, school, district must be chosen for the indicator.

Indicator Name	Description	Formula/Derivation	Periodicity	Data Elements Required
Days Absent	Count of days absent for current school year	Sum number of absences for the school year	Daily	Absent (yes, no) by class
	Count of days absent for prior school year(s)			

Step 3. Set a THRESHOLD.

D = R P = O C = O An = O Au = O

Describe the threshold, benchmark, or objective that defines when action is required and what action is appropriate.

Threshold Name	Description	Value on Indicator
Excessive Absences—Current School Year	Excessive number of absences for the current school year	>2
Excessive Absences—Prior School Year	Excessive number of absences for the prior school year	>9

Step 4. Identify the Required DATA ELEMENTS.

D = R P = R C = R An = R Au = R

DataSpecs contains metadata standards aligned with NCES's Data Handbooks, SIF, EDEN/ED*Facts*, and other national and state standards for education data elements.

Data Element Name	Element Identifier(s)	Definition	Code Set	Standards Linked to Element
Absent	E0001	Code indicating the student was absent	Yes, No	NCES Student Data Handbook (Element 234); SIF (Attendance Object)

Step 5. Identify the DATA COLLECTION that gathers the required data.

D = R C = R P = R An = R Au = R

To ensure that the required data are actually collected, the process for that collection must be determined and documented.

Collection	Collection Identifier	Date(s) & Periodicity	Office Responsible	Data Elements Collected
Teacher/Class/ Student Report	C0001	Every two weeks	Core Data Office	Student ID, Student Name, District, School, Teacher, Class, Absent

Step 6. Identify the DATA REPOSITORY that is the authoritative data source for the required data.

D = R C = R P = R An = R Au = R

To ensure that the required data are available when needed and to determine where the authoritative source of those data exists, the data repository must be identified and documented.

Repository	Repository Identifier	DBA	Table(s)	Data Elements Stored
Data Warehouse	DW01	Sadie Smith	ATTN	Absence

Step 7. Describe the D3M ACTION REPORT to be generated to inform the action, decision, or question, or to satisfy the profile or compliance requirement.

D = R C = R P = R An = R Au = R

The characteristics of the output of the report must be described to ensure that the audiences can understand and interpret the information as easily as possible.

D3M Action Report Name	Media Available	Date(s) & Periodicity	Office Responsible	Process for Accessing the Report--FERPA
Attendance Alert for Students at Risk of Excessive Absences	Alert on Web Portal; E-mail; PDF	Daily	Core Data Office	Web Portal or E-Mail

Step 8. Determine the Action on the indicator that is to be taken at each threshold level.

D = R C = O P = O An = O Au = O

Various values on the indicator will be associated with different actions. Both those values and the associated action must be determined and documented.

Level 1	Level 2	Level 3
Status: Student meets or exceeds threshold of >2 absences in current school year AND threshold of >9 absences in prior school year. Action: Counselor and teacher coordinate contact with parents.	Status: Student meets threshold of >9 absences in prior school year but not >2 absences in current school year. Action: Counselor and teacher coordinate monitoring of student behavior.	Status: Student does not meet prior year threshold. Action: Student is not included in this D3M Action Report. Counselor and teacher monitor attendance for changes.

Step 9. Design the output format for the Action Report.

D = R C = R P = R An = R Au = R

Create the display template for the access of the data. Specify the RDL code for the report and provision it to the report generation application.

Action Report:	**Attendance Alert for Students At Risk of Excessive Absences**			
Date: Wednesday, February 21, 2007	**School:** Jefferson Middle School	**Grade Level:** 7	**Students Listed:** >2 current year absences AND >9 prior year absences	**Action Required:** Parent conference within two days
Student	**MOSIS ID**	**Teacher**	**Current Year Absences**	**Prior Year Absences**
Abrams, Joe	1234567891	B. Wilson	3	12
Zamora, Billy	1234543216	A. Ramos	3	21
Zimmer, Jane	1234565432	K. Clark	3	10

Step 10. Access the Action Report.

D = R C = R P = R An = R Au = R

Ensure that the intended audience has access to the Action Report. The D3M Education Portal™ provides a means for delivering action alerts directly to individuals' monitors.

Attachment B: Decision Question Matrix for Student Performance Measures

Status – single measurement Trend – multiple measures over time

Decision Example 1	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Are we satisfied with the performance of our students?	STATUS: How did the students perform? (None, Self)	Individual	How did the individual perform? (Text)	How did the individual perform? (Descriptive)	Not Applicable	Not Applicable	Not Applicable	What factors explain the results?
		Group	How did the group perform? (Text)	How did the group perform? (Descriptive)	Not Applicable	Not Applicable	Not Applicable	What factors explain the results?
	TREND: How did the students perform? (None, Self)	Individual	How did the individual perform over time? (Text)	How did the individual perform over time? (Text)	Not Applicable	Not Applicable	Not Applicable	What factors explain the results?
		Group	How did the group perform over time? (Text)	How did the group perform over time? (Text)	Not Applicable	Not Applicable	Not Applicable	What factors explain the results?

Decision Example 2	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Which schools should be investigated to determine if they have effective practices to share with others?	STATUS: How did performance compare to the performance of a reference group? (Norm, Group)	Individual	How did the individual perform compared to the reference group? (Text)	How did the individual perform compared to the reference group? (Descriptive)	What was the actual performance difference between the individual and the reference group? (Descriptive)	Was there a statistically significant difference between the individual's performance and the reference group performance?	Was there a practically significant difference between the individual's performance and the reference group performance? (Effect Size)	What factors explain the results?
		Group	How did the group perform compared to the reference group? (Text)	How did the group perform compared to the reference group? (Descriptive)	What was the actual performance difference between this group and the reference group? (Descriptive)	Was there a statistically significant performance difference between this group and the reference group? (ANOVA)	Was there a practically significant performance difference between this group and the reference group? (Effect Size)	What factors explain the results?
	TREND: How did performance compare to the performance of a reference group? (Norm, Group)	Individual	How did the individual's performance over time compare to that of a reference group? (Text)	How did the individual's performance over time compare to that of a reference group? (Text)	What was the actual performance difference between the individual and the reference group over time? (Descriptive)	Was there a statistically significant difference between the individual's performance and the reference group performance over time?	Was there a practically significant difference between the individual's performance and the reference group performance over time? (Effect Size)	What factors explain the results?
		Group	How did the group's performance over time compare to that of a reference group? (Text)	How did the group's performance over time compare to that of a reference group? (Text)	What was the actual performance difference between this group and the reference group over time? (Descriptive)	Was there a statistically significant performance difference between this group and the reference group over time? (ANOVA)	Was there a practically significant performance difference between this group and the reference group over time? (Effect Size)	What factors explain the results?

Decision Example 3	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Which schools failed to meet the annual objectives for No Child Left Behind and require intervention?	STATUS: Were the objectives met? (Goals or Standard)	Individual	Did the individual meet the goal/standard? (Text)	Did the individual meet the goal/standard? (Descriptive)	What was the actual difference between the individual's performance and the goal/standard? (Descriptive)	Was there a statistically significant difference between the individual's performance and the goal/standard?	Was there a practically significant difference between the individual's performance and goal/standard? (Effect Size)	What factors explain the results?
		Group	Did the group meet the goal/standard (Text)	Did the group meet the goal/standard? (Descriptive)	What was the actual difference between the group performance and the goal/standard? (Descriptive)	Was there a statistically significant difference between the group performance and the goal/standard?	Was there a practically significant difference between the group's performance and the goal/standard? (Effect size)	What factors explain the results?
	TREND: Were the objectives met? (Goals or Standard)	Individual	Did the individual improve in meeting the goal/standard over time? (Text)	Did the individual improve in meeting the goal/standard over time? (Text)	What was the actual difference between the individual's performance and the goal/standard over time? (Descriptive)	Was there a statistically significant difference between the individual's performance and the goal/standard over time?	Was there a practically significant difference between the individual's performance and goal/standard over time? (Effect Size)	What factors explain the results?
		Group	Did the group improve in meeting the goal/standard over time? (Text)	Did the group improve in meeting the goal/standard over time? (Text)	What was the actual difference between the group performance and the goal/standard over time? (Descriptive)	Was there a statistically significant difference between the group performance and the goal/standard over time?	Was there a practically significant difference between the group's performance and the goal/standard over time? (Effect size)	What factors explain the results?

Decision Example 4	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Within low-performing schools, which subgroups require focus to meet performance objectives?	STATUS: How did the subgroups perform? (Disaggregated Group)	Individual	How did the student perform compared to others in each subgroup? (Text)	How did the student perform compared to others in each subgroup? (Descriptive)	What was the actual difference between the individual and the subgroups? (Descriptive)	Was the difference between the individual and the subgroups statistically significant?	Was the difference between the individual and the subgroup practically significant? (Effect Size)	What factors explain the results?
		Group	How did the group perform compared to each subgroup? (Text)	How did the group perform compared to each subgroup? (Descriptive)	What is the actual difference between this group and the subgroups? (Descriptive)	Was the difference between this group and the subgroups statistically significant? (ANOVA)	Was the difference between subgroups practically significant? (Effect Size)	What factors explain the results?
	TREND: How did the subgroups perform? (Disaggregated Group)	Individual	Did the individual improve in meeting the goal/standard over time compared to the performance of subgroups? (Text)	Did the individual improve in meeting the goal/standard over time compared to the performance of subgroups? (Text)	What was the actual difference between the individual and the subgroups over time? (Descriptive)	Was the difference between the individual and the subgroups statistically significant over time?	Was the difference between the individual and the subgroup practically significant over time? (Effect Size)	What factors explain the results?
		Group	Did the group improve in meeting the goal/standard over time compared to the performance of subgroups? (Text)	Did the group improve in meeting the goal/standard over time compared to the performance of subgroups? (Text)	What is the actual difference between this group and the subgroups over time? (Descriptive)	Was the difference between this group and the subgroups statistically significant over time? (ANOVA)	Was the difference between subgroups practically significant over time? (Effect Size)	What factors explain the results?

Decision Example 5	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Where do we focus our remediation resources?	STATUS: How did performance across skill areas compare? (Content / Skill Breakouts)	Individual	How did the individual's performance compare across skill areas? (Text)	How did the individual's performance compare across skill areas? (Descriptive)	How did the individual's performance compare across skill areas? (Descriptive)	Was there a statistically significant difference in the individual's performance across skill areas?	Was there a practically significant difference between the individual's performance across skill areas? (Effect Size)	What factors explain the results?
		Group	How did the group's performance compare across skill areas? (Text)	How did the group's performance compare across skill areas? (Descriptive)	How did the group's performance compare across skill areas? (Descriptive)	Was there a statistically significant difference in the group's performance across skill areas? (ANOVA)	Was there a practically significant difference between the group's performance across skill areas? (Effect Size)	What factors explain the results?
	TREND: How did performance across skill areas compare? (Content / Skill Breakouts)	Individual	How did the individual's performance over time compare across skill areas? (Text)	How did the individual's performance over time compare across skill areas? (Text)	How did the individual's performance compare across skill areas over time? (Descriptive)	Was there a statistically significant difference in the individual's performance across skill areas over time?	Was there a practically significant difference between the individual's performance across skill areas over time? (Effect Size)	What factors explain the results?
		Group	How did the group's performance over time compare across skill areas? (Text)	How did the group's performance over time compare across skill areas? (Text)	How did the group's performance compare across skill areas over time? (Descriptive)	Was there a statistically significant difference in the group's performance across skill areas over time? (ANOVA)	Was there a practically significant difference between the group's performance across skill areas over time? (Effect Size)	What factors explain the results?

Decision Example 6	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Which schools perform poorly on multiple measures and require intervention?	STATUS: Did multiple performance measures agree? (Multiple Measures)	Individual	How did the individual's performance compare across multiple measures? Did they agree? (Text)	How did the individual's performance compare across multiple measures? Did they agree? (Descriptive)	Was there an actual difference across multiple measures of the individual's performance? (Descriptive)	Was there a statistically significant correlation across multiple measures of the individual's performance? (Correlation)	Was there a practically significant difference across multiple measures of the individual's performance? (Effect Size)	What factors explain the results?
		Group	How did the group's performance compare across multiple measures? Did they agree? (Text)	How did the group's performance compare across multiple measures? Did they agree? (Descriptive)	Was there an actual difference across multiple measures of the group's performance? (Descriptive)	Was there a statistically significant correlation across multiple measures of the group's performance? (Correlation)	Was there a practically significant difference across multiple measures of the group's performance? (Effect Size)	What factors explain the results?
	TREND: Did multiple performance measures agree? (Multiple Measures)	Individual	How did the individual's performance over time compare across multiple measures? (Text)	How did the individual's performance over time compare across multiple measures? (Text)	Was there an actual difference across multiple measures of the individual's performance over time? (Descriptive)	Was there a statistically significant correlation across multiple measures of the individual's performance over time? (Correlation)	Was there a practically significant difference across multiple measures of the individual's performance over time? (Effect Size)	What factors explain the results?
		Group	How did the group's performance over time compare across multiple measures? (Text)	How did the group's performance over time compare across multiple measures? (Text)	Was there an actual difference across multiple measures of the group's performance over time? (Descriptive)	Was there a statistically significant correlation across multiple measures of the group's performance over time? (Correlation)	Was there a practically significant difference across multiple measures of the group's performance over time? (Effect Size)	What factors explain the results?

Decision Example 7	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Which schools require changes to perform at the level of similar schools?	STATUS: How did similar students perform? (Prior Performance and Student Characteristics)	Individual	How did the student perform compared to predicted performance based upon similar students' performance? (Text)	How did the student perform compared to predicted performance based upon similar students' performance? (Descriptive)	How did the student perform compared to predicted performance based upon similar students' performance? (Regression/HLM)	How did the student perform compared to predicted performance based upon similar students' performance? Was the difference statistically significant? (Regression/HLM)	How did the student perform compared to predicted performance based upon similar students' performance? Was the difference practically significant? (Effect Size)	What factors explain the results?
		Group	How did the group perform compared to predicted performance based upon similar students' performance? (Text)	How did the group perform compared to predicted performance based upon similar students' performance? (Descriptive)	How did the group perform compared to predicted performance based upon similar students' performance? (Regression/HLM)	How did the group perform compared to predicted performance based upon similar students' performance? Was the difference statistically significant? (Regression/HLM)	How did the group perform compared to predicted performance based upon similar students' performance? Was the difference practically significant? (Effect Size)	What factors explain the results?
	TREND: How did similar students perform? (Prior Performance and Student Characteristics)	Individual	How did the student perform over time compared to predicted performance based upon similar students' performance? (Text)	How did the student perform over time compared to predicted performance based upon similar students' performance? (Text)	How did the student perform compared to predicted performance based upon similar students' performance over time? (Regression/HLM)	How did the student perform compared to predicted performance based upon similar students' performance over time? Was the difference statistically significant? (Regression/HLM)	How did the student perform compared to predicted performance based upon similar students' performance over time? Was the difference practically significant? (Effect Size)	What factors explain the results?
		Group	How did the group perform over time compared to predicted performance based upon similar students' performance? (Text)	How did the group perform over time compared to predicted performance based upon similar students' performance? (Text)	How did the group perform compared to predicted performance based upon similar students' performance over time? (Regression/HLM)	How did the group perform compared to predicted performance based upon similar students' performance? Was the difference statistically significant over time? (Regression/HLM)	How did the group perform compared to predicted performance based upon similar students' performance over time? Was the difference practically significant? (Effect Size)	What factors explain the results?

Decision Example 8	Question (Comparison for Interpretation)	Level of Analysis	Narrative	Descriptive Statistics	Actual Difference	Statistical Significance	Effect Size	Analysis
Where should we invest our money to achieve the most learning?	STATUS: What was the cost / benefit? (Cost Per Unit Difference)	Individual	What was spent to achieve the measured performance level of the student? (Text)	What was spent to achieve the measured performance level of the student?	What was spent to achieve the measured performance level of the student?	Not Applicable	Not Applicable	What factors explain the results?
		Group	What did it cost to achieve the measured performance level of the group? (Text)	What did it cost to achieve the measured performance level of the group? (Descriptive)	What did it cost to achieve the measured performance level of the group? (Descriptive)	Not Applicable	Not Applicable	What factors explain the results?
	TREND: What was the cost / benefit? (Cost Per Unit Difference)	Individual	What was spent to achieve the measured performance level of the student over time? (Text)	What was spent to achieve the measured performance level of the student over time? (Text)	What was spent to achieve the measured performance level of the student over time?	Not Applicable	Not Applicable	What factors explain the results?
		Group	What did it cost to achieve the measured performance level of the group over time? (Text)	What did it cost to achieve the measured performance level of the group over time? (Text)	What did it cost to achieve the measured performance level of the group over time? (Descriptive)	Not Applicable	Not Applicable	What factors explain the results?

From Information to Insight – The Point of Indicators

Foreword

By C. Jackson Grayson, Jr., APQC's Founder and Chairman

I like this paper. I like it because it focuses on action. And–perhaps surprising to some–not on data per se. Yes, it is about "data-driven-decision making" (D3M), but its starts with the end in mind…the action an educator might want to make, and works backward to locate, collect, and synthesize the various data, information, knowledge, insights, indicators, and indexes that will help to make the decision.

That's the way D3M should work.

In most schools and districts that Glynn and I are familiar with, that's not the case.

The data are the starting point. It begins with "a circle of inquiry." Typically collaborative teams look at the data, raise issues of access and equity, reflect and perhaps collect or discover new data and knowledge, look for patterns and trends in the data. Questions multiply which leads to smaller, focused ones about particular students, or content. More data may be gathered. Data spurs reflection, sparks dialog, informs professional development. They ask questions, perhaps look for root cause and trends, and move toward a decision. Hence the name "data-driven decision making."

Very logical. Very common. And useful. But, as Glynn believes and argues in this paper, deficient and inefficient in making good use of the data for decisions.

Glynn reminds us that any good researcher knows that you begin a dissertation or any good well-researched project with a clear statement of the question (hypotheses) to be answered. Then, an only then, data is collected and analyzed to prove or disprove the action or hypotheses. Most D3M does the opposite. It gets the data first, and begins the circle of inquiry to reach a decision. What do the data tell us? Bad research. Bad for decision making.

To understand Glynn's model, look at Figure 1. The diagram will lead you through the processes described in the text moving from "data" to "information" to "Indicators" to "Index," and finally to "Insight" to reach the decision. The diagram clearly shows that the paths are not fixed (in the real world or the model), but may move to "Insight" from any one of three levels.

Glynn explains the terms, but they are pretty intuitive. "Data" (such as "counts") become "Information" when the data are organized and presented in a usable format, such as schools ranked in percents of mobile students.

"Indicators" are statistics placed in context for interpretation. Indicators could be schools ranked by mobility rate and change in mobility rate across years, or even opinions based on surveys. There is also a discussion of the value of looking at both

leading and lagging indicators. This needs further discussion, perhaps in his next paper.

Then an "Index" is created, a combination of related "Indicators" weighted to summarize a state or a trend presented on a scale useful for comparisons. Indexes in other settings are the unemployment rate, inflation rate, stock exchange indices, housing starts, or a spider chart of a medical diagnosis.

The final "Insight' occurs when you have considered the data, information, indicators, indexes, and the answer is clear. The doctor says "come back next year," or "report immediately to the emergency room." A decision is made. The decision now becomes clear, almost intuitive, after going through the other steps. It is a decision now based on insight (call it judgment if you wish) after a very careful use of some or all of the previous steps.

Anyone at this point may feel this model is too complicated. One, the real world is complicated. Two, Glynn makes clear the process can be short circuited in several ways, bypassing the hierarchical looking Figure 1 diagram at almost any time.

There are three omissions I'd like to see addressed in future papers.

First, incorporate the use of "process" data. Most of the data in all data warehouses are inputs and outcomes data, not process data. Yet processes determine outcomes data, and if outcomes are to be changed, processes must be changed. It could still fit in his model, but it isn't an explicit part.

Second, use "probabilities" around point estimates of data, information, etc. Making probability estimates around point estimates is one of my private crusades to get into education decision making. Most data are written or calculated as though they are a certainty--that this is the only possible outcome, whether it's graduation rate, mobility rate, or even test data. Everyone makes these point estimates instead of an estimate of the distribution of outcomes around the point estimates. They wriggle out by saying it's a "highly likely" estimate, or "best I can do" estimate—but how high or what's your personal "best"? Is the distribution around the outcome normally distributed or skewed, and what's the dispersion. A further refinement.

Finally, I'd like to point out to Glynn and to readers that it ain't over yet. A decision is not action. A decision is a decision. It's not action. Knowing is not doing. Only doing is doing. Add another box on top called "doing' or "implementation."

Glynn has already started down this point of moving to action previewed with his earlier excellent paper on the same theme: "Actions Speak Louder than Data." His papers are helping to close the 'Knowing-Doing" gap. Keep on going, Glynn!

Introduction

We owe you an apology for wasting your time. The reports you have been receiving are not of great use. The statistics you get about education are abundant but confusing. Even though you have more data than ever before, you are not getting much more than the same numbers broken down for smaller subgroups. The colorful dashboard on your screen looks like a 1949 Plymouth's instruments. If you want to run reports yourself, you must enroll in a three-day course and then be sure to use those skills daily or you'll forget.

Maybe this paper will be the start of a future in which reports will generate insight into the actions you should take. They will be based upon indicators that synthesize data into usable bites. You will save time, but even better, data driven decision making will be worth the effort.

Insight is seeing the solution, realizing what needs to be done.

Our search for actionable data is not a search for data at all, but a quest for the insights we need to inform our actions. Once we know what insights and the associated actions we seek, we must inform them. Identifying and capturing the appropriate data can then be pursued. Aligning all these processes into a system for decision support is the goal of what this paper describes and calls the Extraordinary Insight Model.

ESP Insight
***Insight**: When a decision maker knows what should be done.*

Name three issues that are hot in education accountability today. My three are…

- Multiple Indicators
- Growth Models
- No Child Left Behind

Each of these requires us to take our education data to a higher level. This paper shows how indicators and indexes address all three and many other issues.

Russell Ackoff, a guru of operations research and systems theory, gets credit for organizing the content of our minds into five categories:

1. **Data** – Symbols that represent values or other concepts we need to measure or record.
2. **Information** – Data that are organized or processed to be useful. Information provides answers to *who, what, where*, and *when* questions.
3. **Knowledge** – Data and information that are applied for a particular use. Knowledge answers *how* questions.
4. **Understanding –** Using knowledge to appreciate *why*.
5. **Wisdom** – Evaluated understanding is wisdom.

The first four relate to the past, and only wisdom relates to the future.

The very popular Howard Gardner morphed Ackoff's categories into his own famous quote, "Information is not knowledge, knowledge is not understanding, understanding is not judgment, judgment is not wisdom. If we have no trouble in gaining access to limitless amounts of information, it will only make it more difficult to decide what is worth paying attention to."

Forgiving him for ending his proposition with a preposition, I find all this to still be somewhat esoteric. For me and possibly those millions of other educators who are searching for a way to understand our data, I prefer a very practical schema that deals specifically with education data intended to support decision making. Data driven decision making (D3M) is what I want to bring into focus. I want Gardner to be proud that we have taken his admonition and are deciding to what it is worth paying attention.

ESP Insight
The highest level, Insight, is our judgment of what should be done—what action is needed.

The model I propose is named after our company's motto, *Extraordinary Insight*™. This parallels Ackoff's categories somewhat with the highest level also relating to the future. The highest level, Insight, is our judgment of what should be done—what action is needed. This is my view of how we elevate our data to the heights decision makers demand. That is where readily usable data inform a judgment. However, along the way, we should not be bound to the notion that there is a hierarchy. There is no established path that our data must travel to be fully actualized into an insight. Oh, sorry, that last sentence strays over into Maslow. In the Extraordinary Insight Model, data can be used to form an insight at any level of synthesis. See Figure 1 below.

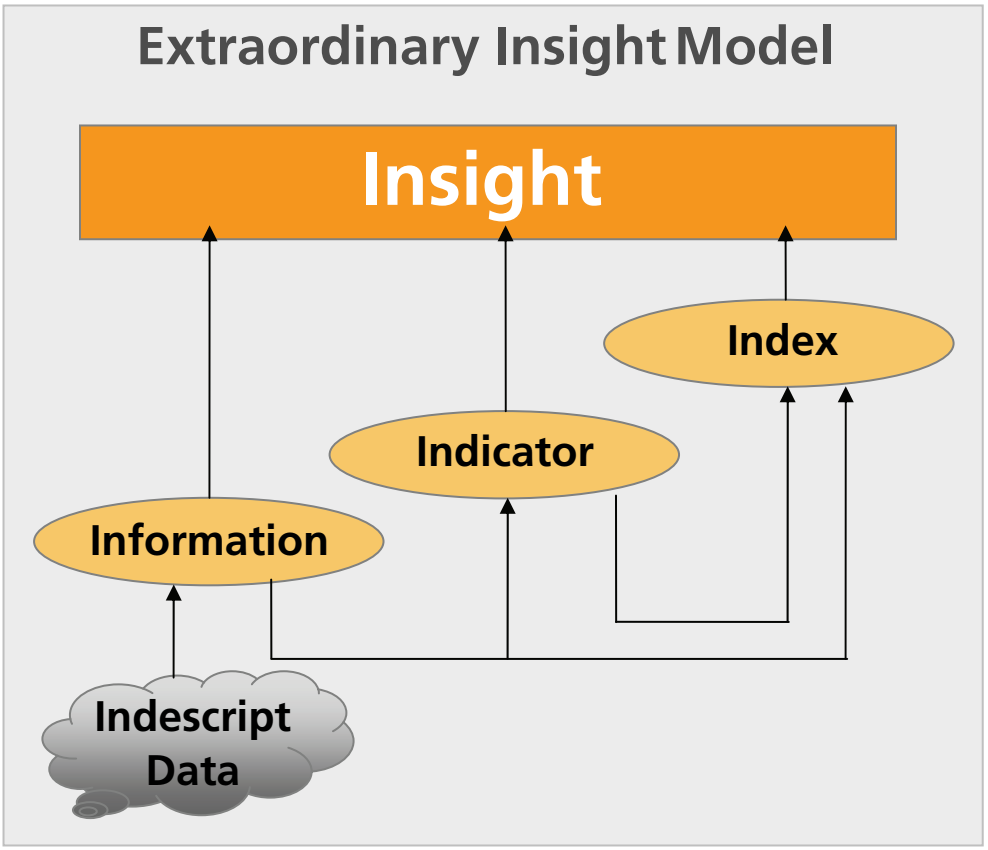


Figure 1: Extraordinary Insight Model

Indesript data must be organized to be usable in this model. That organization turns the data into information. At this point, information can directly inform an insight, or contribute to the derivation of either an indicator or an index. An indicator can either contribute directly to an insight or be combined with other indicators to create an index. The great advantage of indicators and indexes is that they synthesize the data into a summary form that is easier to understand.

The Extraordinary Insight Model

The four categories of data that lead to insights are described below along with an example of each.

1. **Indescript Data**
   Available data that are not focused or organized for use
   - Count of students new to each school

2. **Information**
   Data that are organized and presented in a usable format
   - Schools ranked percent of mobile students

3. **Indicator**
   A statistic placed in context for interpretation
   - Schools ranked by mobility rate and change in mobility rate across years (excluding those matriculating from a lower grade level in a natural assignment pattern)

4. **Index**
   A combination of related indicators weighted to summarize a status or trend; or a single indicator that is presented on a scale useful for comparisons
   - Each school placed on a scale that represents a weighted combination of number of mobile students, percent of mobile students, number of disruptively mobile students (those moving in and out of a school during the school year), and percent of disruptively mobile students.

5. **Insight**
   The relationships, impact, and effect portrayed by the interaction of multiple indicators or indexes that inform a judgment by a decision maker. The insight comes when we decide what action is indicated by the data.
   - Schools with high and growing disruptive mobility rates have lower academic growth measures and require early intervention to prevent the schools from becoming persistently low performing on adequate yearly progress (AYP).

ESP Insight
An insight occurs when we have considered the data (knowledge, indicators, indexes) and the answer is finally clear.

An insight occurs when we have considered the data (knowledge, indicators, indexes) and the answer is finally clear. We apprehend the true nature of a situation. We have almost an intuitive understanding of our problem and its solution. Insight is not data or information at all. Insight is what we discover and conclude from understanding and interpreting all the data available to us.

ESP Insight
I want us to seek indexes that do more to synthesize the data for us than a simple indicator or statistic does.

Differentiating an indicator from an index is a bit difficult. There is not a distinct line between the two, but the distinction is important because I want us to seek indexes that do more to synthesize the data for us than a simple indicator or statistic does. If an indicator does some work for us, then it may be elevated to the status of an index. A truly useful index presents a value on a scale that allows comparisons of both trend and amount. So the categorization of a statistic as an

indicator or an index is somewhat subjective. Here are a few examples of my opinions. (See Figure 2.)

Indicator	Index	Index Components
Student Average Daily Attendance	Opportunity for Quality Instruction Index	-Student Average Daily Attendance -Teacher Attendance -Classes Lost to Non-Academic Activities
Mobility Rate	Disruptive Mobility Index	-Mobility Rate -Mobility During School Year Rate -Total Number of Moves -Total Number of Moves During School Year
Promotion Rate	Pace Toward Graduation Index	-Students Overage for Grade Level -Failing Grades per Grading Period -Failed Courses -Credits Remaining vs. Semesters before Normal Graduation Date
Percent of Certified Teachers	Classes Taught by Highly Qualified Teachers	-Highly Qualified Teacher Criteria Status -Core Classes Taught by Highly Qualified Teachers
Percent Students Proficient	Adequate Yearly Progress	-Percent Students Proficient by Test Area -Percent Students Proficient by Subgroup -Percent Students Participating in Assessments -Percent Students Meeting Alternate Indicators -Number of Students Assessed (Reliability)

Figure 2: Differentiating Indicators and Indexes

Our shared goal in all this is to pull ourselves above the current quality of reports and analyses that is being generated by data warehouses and reporting tools. We need to be at the level of collecting Max Yield Data that have been collected and reported in response to performance on indicators with thresholds that determine actions to take. When we collect too much data (I hear Gardner's voice again), the data we really use competes for resources. Because our indicators rely upon data, we must define them well and make the case that these data are maximum yield to the organization. This sequence and Max Yield Data are described in a prior Optimal Reference Guide from March, 2007, *Actions Speak Louder than Data*, available for download at www.espsg.com/resources.php.

ESP Insight
Max Yield Data: *data that everyone agrees are worth the effort to collect, store, and report.*

What's the Point of Indicators?

The education enterprise is underutilizing indicators for improvement. The reason? I believe educators have not yet managed their data to have the right data, at the right time. If I am correct, then if we redesign our education decision support systems to put the right data in front of educators in a timely manner, then data driven decision making, based upon valid indicators, will occur.

Last month's Optimal Reference Guide on action reports (*Actions Speak Louder than Data*) emphasized the role of an indicator with threshold levels aligned with appropriate actions. This month, we are tackling the indicators themselves.

Indicators give us quick guidance for forming opinions, a.k.a. making decisions that inspire our insights.

Indicators are data points that inform our judgment about the status of an entity (e.g., individual, program, or organization). We rely upon indicators to give us facts—even if those facts are representations of opinions.

ESP Insight
Indicators save us time.

So, what's the point of indicators? Indicators save us time. Simply put, indicators present a clear picture of status or trend. For all of us who are too busy to analyze raw data or detailed reports, indicators are our data shorthand. For those of us who defer to experts to tell us their conclusions, indicators are our data digest headlines.

If this is the point of indicators, then we can understand what describes an excellent indicator—simple, understandable, usable, credible, comparable, available, and valid.

Because an index deals with multiple indicators and how they relate to each other to form a single indicator, indexes must also be simple, understandable, usable, credible, comparable, available, and valid. Being simple for an index means presenting a single scale value, not that the derivation of that value is simple. In fact, an index can be as complex as necessary as long as the result is a single, simple value.

Action Reports

In *Actions Speak Louder than Data*, the process for creating and delivering reports that inform actions was detailed. This paradigm led to the development of a logical process, which is called the D3M Action Report Framework™. There are 10 steps that lead us from the statement of an action to be taken through the production of a report that informs that action. Central to this process are indicators and thresholds.

1. Describe a decision, compliance requirement, profile need, analysis desire, or audit need.

   One of the most essential insights we were taught in our graduate inferential statistics courses was that every experiment must begin with a clear statement of the question to be answered. This is even more essential in the real world because we are too often accepting of available data or statistics rather than independently determining the data we actually need.

2. Define an indicator.

   Indicators give us quick guidance for forming opinions. Indicators are data points that inform our judgment about the status of an entity (e.g., individual, program, or organization). We rely upon indicators to give us facts—even if those facts are representations of opinions from surveys.

ESP Insight
Indicators give us quick guidance for forming opinions.

3. Set a threshold.

   The threshold sets the level on the indicator at which action is needed. There may be multiple thresholds, all indicative of different actions.

ESP Insight
The threshold sets the level on the indicator at which action is needed.

4. Identify the required data elements.

   The indicator and the threshold determine what data are required. The data must be operationally defined in the organization's data dictionary.

5. Identify the data collection that gathers the required data.

   Within an organization, the identified data must be collected to match both the operational definition and the periodicity required to determine levels on the indicator.

6. Identify the data repository that is the authoritative data source for the required data.

   There must be an available, trustworthy, and authoritative source for the required data. To authenticate the quality of the data, the provenance of the data must be established. For those who do not watch the phenomenon called "The Antiques Roadshow," but view instead any one of a dozen legal dramas, the data equivalent of provenance is chain of possession. Who or what system has handled the data from its initial entry through all the exchanges to a final authoritative data store from which an action report will be produced?

7. Describe the Action Report™ to be generated to inform the action.
   To plan for the production of the actual report, several characteristics must be known. These include the report's title, periodicity (schedule), office responsible, media, confidentiality level, etc.

8. Determine the action to be taken at each threshold on the indicator.
   This is the step where the process goes beyond the typical report. This is where an action report is differentiated from a compliance report or profile. Each threshold level that has been identified needs to be aligned with the action that needs to be taken.

9. Design the output format for the report.
   I often caution educators to begin with the report they eventually want to see. However, the prior eight steps are necessary to get to that report ideal. Now is the time to sketch out the actual report to be generated—in the format that supports its use.

10. Access the Action Report.
    This is the action step. The intended user of the report must get it. This can be by "push" (someone sends the report to the user) or by "pull" (the user requests the report). The point of many action reports is to alert the user that something needs to be done, not to expect the user to be aware of action needed. Therefore, pushing an action report to the user should always be the first consideration.

This process assists an organization in collecting data that are valuable—and saving the time and effort to collect less valuable data. Max Yield Data refers to data that everyone agrees are worth the effort to collect, store, and report. (Download the *Max Yield Data* presentation, available at http://www.espsg.com/resources.php.) However, even Max Yield Data must be vetted against their value for informing a decision, determining an action, or answering a question.

One criterion for Max Yield Data is that they are required to inform a decision, specifically a decision that results in action being taken. We already debunked actionable data as a misnomer because data, in their raw form, are seldom interpretable in a decision-making situation. Statistics, metrics, analytics, and indicators are the ingredients of which decisions and actions are made. These derived values are only useful when they are presented in a report that is formatted and available at the moment it is needed and the user is ready to make a decision.

If you are like me, you've seen enough reports. Tables of numbers, often beautifully presented in graphical forms—in living color. However we can be left thinking—so much data, but so little that I would actually use to form a decision. There are several reasons that have arrived at this point. First, we have rushed to put all the data we can find into our data warehouses without a thorough vetting of which data rate the designation of Max Yield Data.

Identifying, Cataloging, and Standardizing Data for Indicators

Underlying indicators have to be standards for our data. ESP Solutions Group uses our data management tool DataSpecs™ to do this for our clients. With this application, we can catalog how data are defined, how they are collected, where they are stored, and where they are used. The Action Report Framework™ guides the use of this application and ensures that the required data elements are available and accessible when an indicator is needed.

ESP Insight
An education agency must have a longitudinal data store.

Because the trends, the growth, the comparisons to benchmarks are all crucial to gaining insights from our indicators, an education agency must have a longitudinal data store. Many of our earlier Optimal Reference Guides discuss the issues related to data warehouses, managing data, data quality, confidentiality, etc. One lesson from best practices within the education information enterprise is that this longitudinal data store (documented over time by an application such as DataSpecs) must maintain and preserve our official statistics. Storing statistics or indicators as they were calculated in the past is a significant commitment by an agency. The business rules change, the systems change, and even the source data change over time. Relying upon the ability to faithfully recalculate statistics in the future is risky. In addition, time is saved in the processing of reports and analyses using stored statistics.

Benchmarking vs. Thresholds

In *Actions Speak Louder than Data*, the term threshold is used but the popular term benchmark is not. Thresholds are defined as those levels on an indicator that determine what action is appropriate. A benchmark is a performance level that has been established either by a standard-setting process or by measuring performance of a comparison group. Therefore, a benchmark is actually one form of a threshold. Benchmarking, the activity, is one method for establishing thresholds. However, I want us to think of a threshold as a point on an indicator that is associated with a specific action. For example, a benchmark of 17% teacher turnover in urban middle schools helps us understand whether or not our own middle school is typical. However, a threshold of 20% teacher turnover being established as the point at which a middle school is required to develop a formal plan for teacher retention is much more definitive—and useful for informing action. The insight that comes from the 20% threshold is that this is the point at which the school can no longer continue as it is. Changes are necessary because 20% has been defined as disruptive turnover.

ESP Insight

A benchmark is actually one form of a threshold. Benchmarking, the activity, is one method for establishing thresholds.

Within our discussion of indicators, a benchmark is a level for comparison not a threshold itself. For excellent advice on benchmarks and benchmarking, I recommend the American Productivity and Quality Center (APQC) as a resource. Their Process Improvement and Implementation in Education (PIIE) project is guiding school districts in the effort to establish process benchmarks for evaluating and improving their productivity. (See www.apqc.org/pile.)

Indicators and Indexes

In economics and finance, an index (for example a price index or a stock market index) is a scale of activity that serves to provide a benchmark of performance—specifically change in performance over time. What is implied in an economic index are the insight and the action. For example, if an inflation index rises, then we should put our investments in Treasury bonds to ride out the coming downturn in the economy.

ESP Insight
The functionality of an index is that users do not need to know the details as long as they trust the index itself.

The functionality of an index is that users do not need to know the details as long as they trust the index itself.

The Bowl Championship Series (BCS) rankings are an index. Multiple measures of a NCAA college football team's ranking are combined and weighted to determine its point on a single scale. An economist/sports fan would track the BCS scale value for a team to determine if over time (either within a season or across seasons) the team is improving or declining. Typical fans would not delve into the inner workings of the BCS rankings, but would focus on their teams' rankings each week. The fan's insight would be whether or not to buy tickets to a particular bowl game or place a wager of a given amount on a favored team.

The BCS ranking is a great example of an index with thresholds and actions associated with each level. The two teams with the highest index values go to the national championship game. After that a complex decision process kicks in where other bowls select among the ranked teams, but some ranked teams must be selected before others. Not many indexes possess such an official and inflexible set of actions associated with values on their scale.

Index = Synthesis of Data (Multiple Indicators)

Now we begin to address the issue of multiple indicators in education. Later on in this paper the Colorado Conundrum illustrates the issue in more detail.

An index is a synthesis of data into a scale. Most of the indexes I have created use multiple indicators with the intent of providing an overall high-level rating. So the distinguishing characteristic between an indicator and an index is synthesis. An indicator is a single statistic; whereas, an index is typically composed of multiple statistics. I have not found this distinction clearly in the literature, but for the purposes of understanding how data can be provided to decision makers, this distinction is very functional. For example:

ESP Insight
The distinguishing characteristic that turns an indicator into an index is synthesis of the data.

> **Indicator**: *Attendance Rate*—the percent of days in membership that students are in class
>
> **Index**: *Opportunity for Quality Instruction Index*—a combination of percent days attended by students, percent days in class by the students' regular teachers, and percent days of class when normal instruction occurs (exclusive of assemblies, special events, and other activities not core to the subject of the class)

The index provides us with an opportunity to represent a higher order concept from our data. In this case, the simple indicator tells us how often students miss class, but the index tells us how often a class period is not maximized for effective instruction with both the student and the regular teacher engaged in core curricular activities.

ESP Insight
As we invent indexes, we can use research, common sense, or a combination of the two.

As we invent indexes, we can use research, common sense, or a combination of the two. The intent is to build a scale that allows us to represent multiple indicators for comparison across time. The comparison could also be to a standard such as an accreditation index with thresholds for ratings.

Formulating an Index

To construct an index, the key is to place each component indicator on a common scale. I like z-scores because they magically transform our data into an equal interval scale with a mean of zero and a standard deviation of 1.0. The magic is that we do not have to meet the assumptions of normal distributions of data because we can force the data into a normal distribution. With different indicators converted to their own z-scores, the overall index can be simply a mathematical combination of all indicators—with or without weighting each indicator for its relative importance. An example is provided in Figure 3.

In North East Middle School, the students' opportunity for quality instruction has been declining. However, the change has been within one standard deviation on the index. The insight here is that, although the school is in a normal range, the trend downward is an alert for action.

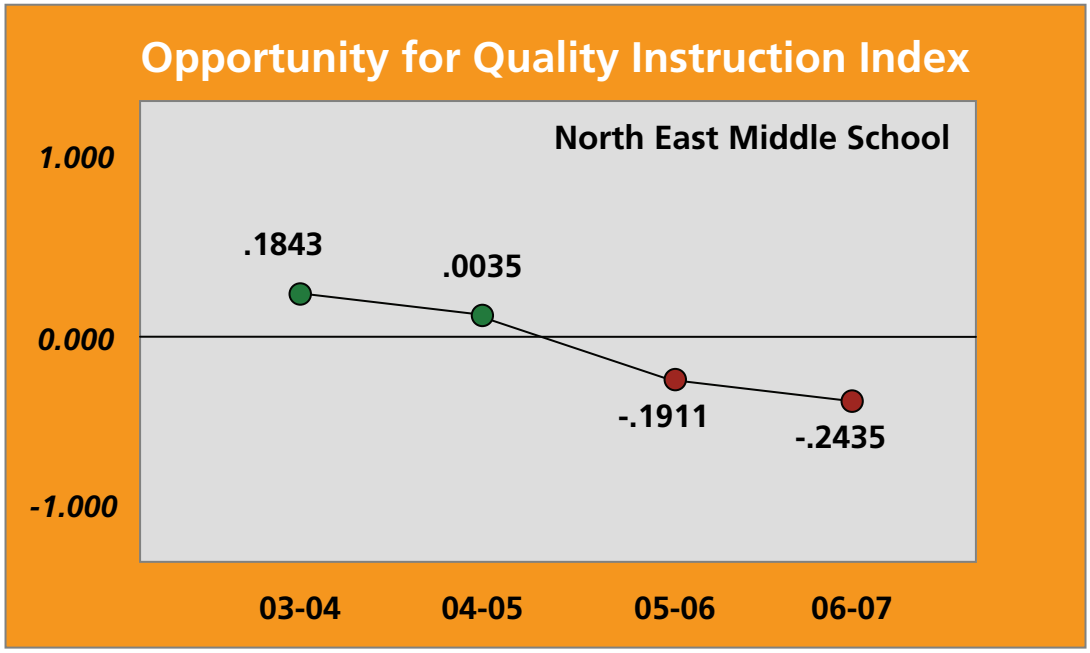


Figure 3: Opportunity for Quality Instruction Index

School:	Index:	Current Index Value:	Trend:	Comment: This school's index value has declined over the past four years from being above average to being below average.
North East Middle School	Opportunity for Quality Instruction	06-07: -.2435	03-04: .1843 04-05: .0035 05-06: -.1911	
Indicator	**School's Z Score (among all middle schools in the state)**	**Weight (determined by district advisory committee)**	**Weighted Value (mean = 0.0; standard deviation = 1.0)**	**Explanation**
Average Daily Attendance by Students (= 93%)	0.808	.50	0.0404	The student's own attendance is key to learning regardless of the activity occurring each day.
Percent Attendance by Regular Teachers (= 91.2%)	-.15	.25	-0.0375	When the regular teacher is present, students learn more and instruction is consistent with the courses scope, sequence, and status.
Percent Days of Normal Instruction (= 90.7%)	-0.1090	.25	-0.02725	Even when the student and regular teacher are present, if non-core activities are occurring, there is a lost opportunity to learn.

Figure 4: Opportunity for Quality Instruction Index

Leading indicators are those that forecast changes or alert us to changes at an early point in time. The Consumer Price Index is a leading indicator of future inflation or even stock market trends. (I must admit having trouble determining for education what a leading indicator is. Then I realized that an indicator may be both a leading and a trailing indicator at the same time. Continue reading.)

Leading indicators in education may be changes in student mobility, English language learners, enrollment in high-level courses, highly qualified teachers, promotion rate, average daily attendance, teacher transfer/turnover rate, diagnostic assessment results, birth rate/housing starts, etc.

Trailing indicators are typically the ones we are most interested in as consumers of education services. They are often the outcome measures in which we are most invested. Assessment scores, graduation rates, schools in need of improvement,

ESP Insight

Our favorite indicators are trailing ones, so we are always looking at history rather than anticipating what we need to be doing right now.

AYP statuses, accreditation ratings, actual enrollments, and actual class sizes may be trailing indicators.

The criticism of education indicators by business experts has been that our favorite indicators are trailing ones, so we are always looking at history rather than anticipating what we need to be doing right now. I think business experts underestimate the expertise of educators. We just have done a better job of formally measuring and reporting our trailing indicators. In fact, educators have an informal, common sense approach to leading indicators that is evident when one talks to counselors, principals, and program managers. They watch for signs of distress in schools and classrooms even if they don't have formal statistical measures and reports to use. These leading indicators can be formalized and improved considerably.

Jack Grayson and I had a lengthy discussion about leading and trailing indicators for education. He tags most indicators as trailing because they are based upon measures of things that have already occurred. As we reviewed the list of indicators in Attachment A, we could see how each might be either leading or trailing. They seem to be currently trailing indicators based upon how we have traditionally used them—late, as a look back on how we did.

So we tried to define a leading indicator. Our best judgment was that if action is the key result of the indicator, then it is a leading indicator. If the indicator measures the result of an action, it is a trailing indicator. Another perspective is that if the indicator forecasts future outcomes, then it is leading. Combining the two then, we concluded that the best leading indicators are those that forecast a future outcome and as a consequence provide us the opportunity to take action that will influence that future outcome positively.

ESP Insight
The best leading indicators are those that forecast a future outcome and as a consequence provide us the opportunity to take action that will influence that future outcome positively.

Insight is when we realize what action can be taken to influence future outcomes. Data driven decision making is this entire process.

Efficiency Indicators – Education has not formalized the use of efficiency indicators. Jack Grayson has urged this change as the most promising way to improve schools. What are his four efficiency indicators?

- Cost Effectiveness—The dollars spent divided by a measure of the outcomes or benefits achieved
- Process Efficiency—Outputs divided by inputs
- Staff Efficiency—Outcomes divided by staff full time equivalents
- Cycle Time—How long a process takes to complete

Ask administrators and managers in education agencies about process or efficiency indicators, and they can talk about the ways they monitor their areas of responsibility. Almost all do this informally. The PIIE project is providing a mechanism for those managers to share benchmarks to formalize the evaluation and improvement of processes.

Converting Trailing Indicators to Leading Indicators

Maybe the problem in education with our focus on trailing indicators is when we measure and report them. Take student mobility for example. If this indicator is reported in an annual profile that is published six months after the end of a school year, that's a real trailing indicator. However, if student mobility is reported on a weekly periodicity, it can be an alert for support for schools experiencing an increased burden for managing enrollment records, staffing changes, and diagnosing new students' needs.

Vital Signs

Being the chief information officer for an urban district with a frenetic superintendent taught me how to react quickly to spur-of-the-moment ideas. After our superintendent recovered from a near heart attack, he visited the Mayo Clinic and was impressed with the manner in which the doctors presented him a single graphical overview of his health condition. One of the aspects of this overview that he particularly liked was that every indicator was displayed on the same scale with lines and colors used to show relative strength across them. He wanted to know why education didn't have a similar technique. From that episode, we designed and published for the next two years Vital Signs. The beauty of Vital Signs was that a glance would tell the observer the general upward or downward trend of the school system over the past six weeks. Comparisons to a five-year historical level and to the previous six-week period were presented.

ESP Insight
Simplicity communicates.

Two lessons learned from this experience are:

1. Simplicity communicates.
2. Audiences for indexes appreciate graphic presentations that provide a quick impression.

ESP Insight
Audiences for indexes appreciate graphic presentations that provide a quick impression.

Education data are more complex and inter-related than data are in other enterprises. Consider the relative simplicity of a medical chart of a patient's vital statistics compared to education statistics. (See Figure 4.)

Health Care Indicator		Education Indicator	
Heart Rate	Count of the number of heartbeats in one minute	Attendance Rate	Count of the number of days attended divided by the number of days of membership within a predetermined time period, with rules for counting a number of tardies as a single absence and using a maximum number of consecutive absences before a student is withdrawn and no longer counted absent; measured at 10 a.m. or at the beginning of second period; excluding expulsions; excluding approved athletic events, field trips, and other co-curricular activities; possibly including excused absences, but not necessarily.

Figure 4: Heart Rate vs. Attendance Rate

A medical professional would argue that I picked their simplest metric, but the reality is that when we go to the doctor's office, our heart rate is measured and recorded—no fancy rules or calculations. In contrast, every state has definitions of attendance and membership with rules for when students get counted in each. To exacerbate education's challenge, those state rules then are interpreted and followed idiosyncratically by school districts, schools, teachers, and attendance clerks—and gamed cleverly by students. That's all before we even try to enter attendance data correctly into a computer and perform the arithmetic accurately.

ESP Insight
States, districts, and schools are working hard to improve data standards and quality.

Don't give up on education statistics. States, districts, and schools are working hard to improve data standards and quality. Also be reassured that for the indicators we are discussing in this paper, timeliness may be more important than complete accuracy.

In 1994, I wrote a paper for the American Educational Research Association on indicators. (***What Dow Jones Can Teach Us: Standardizing Education Statistics and Indicators***) At that time, there were seven characteristics of education indicators systems that were evident. Those are repeated below with an updated perspective on each.

1. Indicator systems too often mold themselves to what is available rather than what should really be collected and reported. How an indicator is calculated is usually determined by what data are available (e.g., a membership count rather than a cumulative enrollment, the arithmetic difference between ninth- and twelfth-grade enrollments rather than an actual count of dropouts).

> ***Today:*** *Still too true. In last month's paper on action reports, the necessity of determining what data are really required to inform an action was emphasized.*

2. Indicator systems can grow to have so many components that audiences have a difficult time sorting them out and drawing a conclusion from them. The complexity is multiplied when statistics are disaggregated by a large number of groups.

> ***Today:*** *The No Child Left Behind Act , to its credit, requires a single Adequate Yearly Progress rating for a school, district, and state. However, it also requires an annual report card with a multitude of indicators—disaggregated by subgroups.*

3. Definitions of indicators across schools and school systems are often too varied to allow reliable comparisons to be made. Variations across locations in the definitions of data elements, the timing of collection, and the accuracy of the reporting cast some doubts on the reliability of some indicators—especially as data are aggregated at the state and national levels.

> ***Today:*** *The U.S. Department of Education (USED) recognized this issue and has recently launched the Education Data Exchange Network (EDEN) to collect their data from states in a more standardized process with improved standards for data definitions. EDFacts is the public reporting application for these data. Individual states have also recognized the need for standardization of data definitions. ESP Solutions Group has worked with USED and over half the states to document and align their data standards.*

4. Data quality is usually unknown or accepted as the best that is available.

Today: *The establishment of data standards described in #3 helps tremendously. However, as our company works with individual states and districts, we are continually confronted with concerns by the educators over data quality. They understand the challenge to produce quality data and the implications of being compared to other entities with suspect data.*

5. The typical evolution of an indicator system is that each indicator reported raises questions that must answered by another indicator, then each indicator must be disaggregated to create sub-indicators, finally there are so many indicators that there is a call for a single indicator or for the agreement upon a few as the most important.

Today: *This will always be an issue. We need to accept the fact that an indicator is not intended to answer all the questions—merely to give us a heads up that we can either accept or seek more information.*

ESP Insight
An indicator is not intended to answer all the questions—merely to give us a heads up that we can either accept or seek more information.

6. Some school systems, through strategic planning efforts, have identified their targeted outcome indicators and have begun to differentiate among those indicators that are descriptions of process, resources, or implementation rather than outcomes.

Today: *Confusion still persists over what is an outcome. For an education agency, outcomes are student performance measures, e.g., academic performance measures, graduation rates, etc. Attendance, discipline, and other measures are process indicators in my book.*

7. Indicator systems are labor intensive. This helps explain why most school systems rely upon the state education agency (SEA) to develop and maintain them.

Today: *Much has changed. Defining, gathering, analyzing, and managing the data are still labor intensive even with automated processes. However, decision support systems have made producing the reports, e.g., web pages, almost too easy. Too easy because more thought and planning are needed to get beyond the typical counts on most web pages.*

What are the Indicators?

- What is the quintessential indicator for education organizations?
  - Graduation from high school
    - In the past, we focused on the negative—dropouts. Then there was a call to be more positive and NCLB asked for graduation rate.

ESP Insight
Graduation from high school is the quintessential indicator.

- What is the penultimate indicator?
  - Enrollment in higher education
    - After successfully launching students from high school, the secondary goal is to see them enroll in higher education.

- What is the societal indicator?
    - Graduates who are productive citizens
        - To be a productive citizen, we have determined that having functional literacy, basic mathematics skills, an understanding of government, and a broad sense of science are needed. We measure those with our academic assessments and course credits. We could also count registered voters and actual voters among our graduates.

For the sake of discussion, Attachment A lists some suggested indicators and indexes.

No Child Left Behind — AYP Index

The No Child Left Behind Act of 2001 has made a great contribution by focusing our attention on significant accountability issues. The Act's shortcomings are well documented. One of the lightning rod provisions requires all subgroups of students to meet every annual objective for a school to make adequate yearly progress. Even though this provision is central to the name of the Act and its intent, the practicality has been that educators are dealing with a large number of schools in need of improvement.

An index would address this issue very well. Robert Linn, former President of the American Educational Research Association, proposes a compensatory model. With a compensatory approach, high achievement that is above the goal in one content area can be used to compensate for achievement that falls below the goal in another area. To accept this model, we must abandon the basic premise of No Child Left Behind that every student must be proficient in every core content area for a school to be acceptable. However, I am comfortable in defining a student as proficient if the student averages proficiency across all areas. I am not comfortable with averaging subgroup performances across areas because that again allows high-performing subgroups to mask the poor results a school is achieving with their low-performing subgroups.

ESP Insight
Averaging subgroup performances across areas allows high-performing subgroups to mask the poor results a school is achieving with their low-performing subgroups.

What might an index for AYP include? First, there should be one index for individual students, then an index to determine AYP for a school.

The Student AYP Index could be simply a weighted sum of scale scores across all areas included in the AYP process. (See Figure 5. Assume a vertical, equal interval scale.)

The school (or district or state) AYP Index could be a weighted sum across proficiency rates on the assessments and alternate indicators. (See Figure 6.) However, this compensator model should be supplemented by continued reporting of the status of every subgroup.

These are greatly simplified examples. We would certainly find many ways to make this more complex. However, that complexity is more of a question of **politimetrics** than one of design. Politimetrics by the way is the setting of our metrics for accountability through a political process. In education, we call this collaboration, involvement of stakeholders, and eventually regulation by a board.

ESP Insight
***Politimetrics** is the setting of our metrics for accountability through a political process.*

Student AYP Index	North East Middle School	Student Identifier: 1234567891	Grade Level: 7
Indicator	**Student's Scale Score**	**Weight (determined by state NCLB plan)**	**Weighted Value**
Mathematics Proficiency	345	35	120.75
Reading, Language Arts Proficiency	367	35	128.45
Science Proficiency	423	30	126.90
Student's Total Compensatory Proficiency Score			376.1
Criterion for Proficiency			350.00
Student's Status	**Proficient**		

Figure 5: Student AYP Index

School AYP Index	North East Middle School		
Indicator	**School's Performance Level**	**Weight (determined by state NCLB plan)**	**Weighted Value**
Percent Students Proficient by Compensatory Scale	84.5%	75	63.38
Graduation Rate	88.7%	25	22.18
Composite AYP Scale Score	85.56		
Annual Objective	85.00		
AYP Status	**MET**		

Figure 6: School AYP Index

Reliability

This would be the opportunity for states to move from their current methods for determining reliability (sampling error) to one grounded in the basic principle of the No Child Left Behind Act (measurement error). This distinction is detailed in an earlier Optimal Reference Guide, *Confidentiality and Reliability Rules for Reporting Education Data*, available for download at www.espsg.com/resources.php. Because the School AYP Index is based upon counts/percents/proportions, nonparametric significance tests with consideration of the standard error of measurement (SEM) for each assessment are appropriate. James Popham and I have discussed these issues since the enactment of the No Child Left Behind Act. We agree that sampling assumptions do not fit how schools get their students—they are not randomly assigned each year to schools. Popham wants states to use a test-retest SEM, which would be larger than the typical internal consistency SEM states use. However, practically, state assessment programs do not invest the dollars or tolerate the burden on schools to determine a true test-retest SEM for their assessments.

ESP Insight

Sampling assumptions do not fit how schools get their students—they are not randomly assigned each year to schools.

Multiple Indicators

This is one of my favorite issues in education. After the wave of assessments for accountability in the 80's and 90's, there was a hue and cry for use of multiple indicators rather than a single test score. The No Child Left Behind Act, despite being characterized as relying upon a single test score, actually requires alternative indicators. The best illustration of this is what I refer to as the "Colorado Conundrum." Their legislature passed separate laws requiring accreditation and annual school accountability reports. When NCLB arrived, AYP became the third major accountability system. Some local districts even are independent enough to maintain their own accountability systems for their schools. So predictably, professional and political groups that often view the world differently joined together to champion an effort to create a single, unified system that uses multiple indicators rather than relying only on their state assessment, the CSAP. In one of their early meetings, there was great passion for measures such as teacher ratings of student progress (after all, as the argument goes, who better knows how well a student is performing than the teacher?) Never mind that the reason accountability systems arose in the first place was the subjective, unreliable, non-comparable nature of teacher judgments, e.g., report card grades. State legislatures determined long ago that there was not only a conflict of interests if teachers provided their own accountability but that the bias may not favor the welfare of the students in some cases. The list of possible indicators to augment CSAP (or in the minds of many there to replace CSAP) began to fill flip chart pages covering the walls. Inevitably, someone pointed out that what they would need is a way to organize and combine all of these multiple measures into a single one that parents could understand as representative of the gestalt across the possibly confusing array of multiple indicators. Then as the speaker pointed out, Colorado would have achieved a single accountability system based upon a single metric.

ESP Insight
In the end what people really want is a single measure—not multiple indicators that present a confusing array of information.

Well, that was a sobering moment. Maybe we had just realized why the three discounted accountability systems had focused on CSAP. In the end what people really want is a single measure—not multiple indicators that present a confusing array of information, possibly conflicting, and require us as individuals to make our own judgment of a school's effectiveness. Conundrum.

The sense throughout the room was that all those nominations for multiple indicators shared a common softness. They would be difficult to standardize for reliability across teachers, schools, and districts. They also tended to stray a bit from being outcome and performance measures. In fact, I came away from the meeting with a renewed appreciation for standardized tests that have been aligned with academic standards and administered following a structured protocol.

ESP Insight
What people want is for their single indicator to be influenced by multiple measures.

The Colorado Conundrum is really not that difficult to analyze. Only on the surface are there inconsistencies. What people want is for their single indicator to be influenced by multiple measures.

Accountability vs. Diagnostics

Sometimes we try to create a single indicator when more than one is needed. The persistent criticism we hear of state assessments is that they do not provide teachers with the information they really need. Well quite frankly, those assessments were not originally intended to do that, nor are they very good at that.

This issue, which I find resistant to logic, is what I call the "Texas Two Step." Texas gets their name in the title simply because they have been dancing around the issue longer than most everybody else. The Texas Two Step is trying to take two steps in opposite directions at the same time. Here's how it goes—playing the Cotton-Eyed Joe now is purely optional.

Legislatures want accountability. Teachers want diagnostics. Psychometricians know that a single assessment cannot do a credible job of both at the same time. However, states persist in trying to develop criterion-referenced tests to give annually for accountability ***AND*** provide useful diagnostic information for teachers. That's just not going to happen. (Yes, I capitalized, italicized, bolded, and underlined ***AND*** to illustrate my frustration.) Consider the facts in Figure 7.

ESP Insight
Legislatures want accountability. Teachers want diagnostics.

Assessment Characteristic	Diagnostic Assessments	Accountability Assessments
Indicator/Index	Proficiency by individual students on standards for knowledge and skills	Rating on an accountability system by level of overall performance
Purpose	Determine student's current level of proficiency so instruction can be targeted at specific needs	Determine whether a student, school, district, or state is performing at a target level
Audience	Teachers and other educators who plan and deliver instructional activities	Public, governmental, administrative, and parental individuals and groups who are stakeholders in the performance of schools and the students they teach
Content	Selected standards that are timely for planning instruction	Broad sampling of content to represent all standards
Number of Items	Varies dependent upon the number of standards being measured	Each general content area may have 35 to 50 items depending upon the time required to respond to each.
Type of Items	Constructed response with multiple choice as appropriate	Multiple choice to maximize objective, quick scoring; constructed response as necessary
Item Selection	As many items as possible for each individual standard being assessed	Small number of items for each of a larger number of standards
Reliability	The larger the number of items for each standard the higher the reliability of the score for each standard	Because measuring individual standards is less important, fewer items over more standards is desired
Timing	On demand as close to the delivery of instruction as possible; not scheduled for all students at the same time	Annually or at selected times; may be scheduled for all students at the same time
Security	Validity and timeliness are more important	Highest security is important
Reuse of Items	Reuse of items for different students is preferred	Reuse of items is problematic because of security concerns

Figure 7: Diagnostic vs. Accountability Assessments

ESP Insight

Develop one testing system that is really good for accountability and another one that is good for diagnosis.

The solution to the Texas Two-Step has been known for decades. Have two testing systems. Develop one that is really good for accountability and another one that is good for diagnosis. I believe that two systems would not cost any more than what states are spending now for one hybrid.

Growth Models

Growth models are inspiring and challenging educators today. The allure is that we will find those schools that are actually very effective even though their students are scoring very low on state assessments. I share that expectation because back in the 80's, we produced a local report on the differences between a regression-based prediction of achievement and actual achievement for all of our schools in Austin. Indeed there were schools that produced gains in excess of those predicted by their students' starting levels of achievement and demographics. From those results, I can say comfortably that a growth model will not find very many schools that would be considered effective after being designated as low-performing by a status indicator. That is unless the growth model adjusts for income.

There is not enough space in this paper to analyze the relative merits of value-added growth models that use regression or hierarchical linear models to create a measure of performance that basically sets a lower standard for low-income or low-performing students than for their high-income or high-performing peers. I prefer a growth index that answers the No Child Left Behind-style question of whether or not a school is moving students along at a pace that will at some point in time elevate them to a proficiency standard.

The single point to be made here is that a growth model should be an index. Regardless of the component indicators and formulas used, the growth for a student or a school can be represented on a scale with thresholds that designate significant gain, unreliable gain, no gain, unreliable loss, and significant loss.

ESP Insight
A growth model should be an index.

An Academic Growth Index

This index is presented merely as an example—not necessarily as a recommendation. (See Figure 8.)

Student Academic Growth Index	**North East Middle School**			**Student Identifier:** 1234567891	**Grade Level:** 7
Indicator	Student's Scale Score in BASE YEAR	Student's Scale Score in CURRENT YEAR	Growth	Weight (determined by state NCLB plan)	Weighted Growth Value
Mathematics Proficiency	345	567	222	35	120.75
Reading, Language Arts Proficiency	367	587	220	35	128.45
Science Proficiency	423	523	200	30	126.90
Student's Total Growth Score	Actual Growth				376.1
	SEM Reliability Range				25.0
	Minimum Reliable Growth				351.1
Criterion for Normal Growth	*Criterion can be a set standard, a projected level based upon Regression or HLM, a calculation of normal growth for a student at the beginning proficiency level, etc.*				350.00
Student's Status	**Met Growth Target—Reliable**				

Figure 8: Student Academic Growth Index

Not a Fan of Newsweek's Rankings

Newsweek has published the last couple of years a ranking of the nation's high schools. This has to be the worst example of an index. Their ranks are based solely on the ratio of Advanced Placement and International Baccalaureate exams taken to the number of seniors. How easy is that to criticize? The author's response to criticism has been simply these are the best numbers he found to use, and they make sense to him. He didn't try hard enough. An index and the rankings that result are serious. I would much prefer to see educators create and publish meaningful indexes than to continue to see magazine writers grab whatever is available and print millions of copies.

ESP Insight

Newsweek's ranking of high schools is the worst example of an index.

Indicator vs. Public Index

ESP Insight
A public index is a very general-level scale intended for a broad audience.

In the 1994 AERA paper, I used the term public index as distinct from an indicator. The distinction between an indicator and a public index is that a public index is a very general-level scale intended for a broad audience. A public index meets these criteria:

1. The audience does not have to know how it is formulated, because its primary purpose is to communicate an otherwise complex and difficult-to-comprehend phenomenon to a lay audience.

2. A relationship across time or to a target is represented.

3. A predictable periodicity of reporting allows the audiences to maintain an impression of both status and trend.

4. There may be multiple components of a public index. Multiple indicators may be combined into a single public index.

Think about the most frequently reported indexes that you can name, for example:

- Consumer Price Index
- Cost of Living Index
- Inflation Index

What we realize is that an important characteristic these share is that they are frequently reported. They are frequently reported because they show changes during the reporting intervals, so there is "news" to report. In addition, they are of interest to us, and we each have an emotional (e.g., optimistic/pessimistic) reaction to them. The periodicity of these indexes is important, because the interpretation of the index is usually tied to the change from some point in the past.

The Dow Jones Average, the index of the financial worth of corporations, is widely used as an indicator of the trends in corporate America. Almost anyone on the street, not just Wall Street, but Main Street, USA, has a feel for what is happening in the financial markets upon hearing phrases such as, "the Dow closed today at its highest point in three weeks with stocks averaging...." Does the average person really understand what the Dow means or how it is calculated, or even what the professionals within the financial markets really think about it? No, but the average person maintains a sense of up or down, good or bad, boom or bust for the whole country based upon the tidbits of radio, television, and newspaper information seen daily about the Dow.

ESP Insight
Education has been thought by some to be too complex to be represented by a single index.

Education has been thought by some to be too complex to be represented by a single index. Some also believe that local education agencies (LEA) or state education agencies (SEAs) differ too much to be compared on a single scale. I prefer to approach these issues with the perspective that public education will never have the confidence of the general public until there is some common sense, simple, frequently reported index of whether or not our students are doing well in school.

One of the first lessons that we can learn from the Dow Jones Average is that the

professionals in the financial industry understand that average well enough to interpret it with great caution. In fact the analogy to an index for schools is excellent in the sense that anyone interpreting the Dow or an education index should ask questions about recent events that could have affected the index, seek more detailed information to assess an individual stock or school, or consider individual components of the index that could be having a temporary undue influence.

The pubic is frustrated that education does not have frequent and generally available indexes that they can follow informally or use to compare their local school to the national trends. They should be frustrated, and Congress was also frustrated when it passed the Hawkins-Stafford Education Improvement Amendments in 1989 and the No Child Left Behind Act in 2001. From those and other laws, efforts have been made to set standards for education information (Standards for Education Data Collection and Reporting, SEDCAR), to define the data elements that should exist in an educational reporting system (National Center for Education Statistics Student Data Handbook), and even national standards for the exchange of student records electronically across computer networks (SPEEDE/ExPRESS, SIF E-Transcript Object, PESC High School Record, and Texas's TREx). The National Education Goals Panel defined the measurement of six goals that came from the Education Summit of 1989. These and other efforts focused on a similar target—to establish within the education industry some standards for collecting and reporting the data required to monitor and manage public education.

ESP Insight

The public is frustrated that education does not have frequent and generally available indexes that they can follow informally or use to compare their local schools to the national trends.

The taxonomy of indexes in that 1994 paper still work well for conceptualizing some important dimensions of a public index. If elementary and secondary education is to gain the public's confidence that we have a finger on the pulse of our schools, we cannot overwhelm them with too many indexes.

Taxonomy of Indicators

Statistics/indicators can follow several periodicities:

- **Long Term**
  - Longer than a year
    - NAEP, OCR Reports
- **Annual**
  - The same time every year
    - So much revolves around the "school year." So many of our indicators are measured once a school year. AYP
- **Periodic**
  - At even intervals shorter than a year
    - Student report cards, meals served
- **On Demand**
  - *Ad hoc* measures made or reported upon demand
    - Recruiting pipeline, tax receipts
- **As Available**
  - Whenever possible
    - Evaluation study findings

ESP Insight
"I believe that process management is the single best way to transform public K-12 education."
C. Jackson Grayson

A second important dimension for indicators is the level being measured:

- Outcome: Accountability level; were the goals and objectives met?
- Process: Implementation level; were the services delivered?
- Resources: Support level; what funds, personnel, equipment, and other materials were applied?
- Context: Pre-existing conditions; what advantages or disadvantages existed that might impact resources, processes, or outcomes?

The best, most useful indicators are those that are available on demand and are up-to-date whenever reported. These can become our leading indicators. Student demographics and course enrollments are two examples. Periodic indicators are the next most useful because they can track changes in time spans of less than a year. Report card grades and attendance reports are two examples.

For the general public, the most useful index would be one that is available periodically, to allow for tracking changes across relatively short time spans; and targeted at the outcome level, to give an overall impression of learning levels. Indicators that describe processes, resources, and context would primarily be for educators and members of leadership teams advising schools.

Possible Public Indexes for Education

Imagine what it would be like to open the morning paper and read, "The National Education Goals Index rose seven points last week to a record high."

Two basic questions are asked by the general public:

1. How many (what percentage) of our students are successfully completing high school (graduating)?
2. Do those graduates have the skills they need?

Although these questions appear to be directed at the end of the public education process, a functional index should be applicable to every level from prekindergarten through grade 12. Thus, the two indexes described here apply to all grades.

Pace Toward Graduation

The first public index is pace toward graduation, the PACE Index (Pace toward Achieving Completion of an Education). We need to define some components of a graduation rate that can be measured across all grades and create from those an index of pace toward graduation. Pace toward graduation would have the distinct advantage of being known for every student at any point in time, as opposed to a graduation rate that can be known only for students at the top end of the school system.

ESP Insight
Pace toward graduation would have the distinct advantage of being known for every student at any point in time.

The concept is that at each age level, a student is compared to other similar students who followed the same pace through the grades. Then the graduation rate for those comparison students becomes this student's PACE Index value. For example, a student who is age 15, with 10 high school credits, and classified as a sophomore would be given a value that was empirically determined by the graduation rate of the most recent group of students to graduate who had these same characteristics at age 15. Let us say that of the class of '06, 78% of the students who these same characteristics at age 15 had actually graduated. This student would then receive a value of .78. Now it would be possible to average these values across all students in a school system to determine the aggregate pace toward graduation of the student population. If that average is .90, then we would estimate that the school system is moving students along at a pace that would predict that the eventual graduation rate would be 90%.

Student Mobility

Student mobility is a prevalent factor in the challenge schools face. Schools constantly inherit students who may be far behind those who have been "native" to the school since the earliest possible grade level. Therefore, there also needs to be an adjustment to the pace toward graduation index to factor out the advantage or disadvantage a school might accrue from mobility. One possibility is that each entering student would receive an adjustment equal to the difference between that student's PACE index value and 1.00. The adjustment would follow the student through the school system. For example, a new or transfer student with an index value of .81 would be assigned an adjustment of +.19 to be added to his/her actual PACE Index value each time it is calculated. This would bring all initial adjusted

values to 1.00. In other words, a school or school system would be held accountable for keeping students on the same pace as they were on upon entry.

This same adjustment could be made for entering preschool and kindergarten students in order to "level the playing field" for all schools. If this were to be done for all students upon entry to a school, then the school's PACE index would be compared to 1.00 or 100%. Meaning if a school's PACE index falls below 1.00, then the students are falling behind the pace predicted.

So for each school or school system, there could be an overall PACE index value and an adjusted PACE index value. The overall value would "predict" based upon current status of the student population what percentage of them will graduate. The adjusted PACE index value would indicate whether or not the students are ahead of or behind the pace "predicted' when they first entered the school or school system.

The model that establishes the index values for each group of students could be developed on a local, state, or national level.

If our goal were to be a 90% graduation rate, when the PACE index reaches a value of .90 or 90%, then the goal could be considered as met, without waiting years to see the actual graduation rate, or using a single graduating class as the basis for measuring the goal.

Separate values could be calculated for each grade level in a school to monitor where students are gaining or losing on the pace of other students.
This type of index could prove to be less vulnerable to the problems associated with other measures. Indeed this measure is filled with subjective criteria for promotion and retention, is greatly influenced by local standards for earning course credits, and is highly dependent upon whether educators are socially promoting students. However, this measure is fundamental to public education; fundamental to the mission of schools: given all the local standards and requirements to which all students are held accountable, are students progressing at a pace that predicts they will graduate rather than drop out?

ESP Insight
Being overage for grade is a strong determinant of graduation potential.

Based upon research by public school systems into the factors that predict dropping out or graduating, the primary predictor that would be expected to play a role in this index would be the students' age – being over age for grade is a strong determinant of graduation potential. Other factors might be more grade specific, such as performance level on a kindergarten readiness assessment, number of failing grades each six weeks in grades 1 -12, and number of credits earned in high school.

Achievement Index

The second index is the Achievement Index. Yes, we could use a national assessment to measure these goals, answer this basic question, and calculate this index. NAEP would be acceptable; selecting one of the currently available, nationally normed achievement tests would be functional; developing a shorter, more general test would work, or performing an equating study across all state assessments. Maybe someone will figure out how to equate all the states' assessments. The instrument is not the primary focus here. In fact, the instrument

ESP Insight
Maybe someone will figure out how to equate all the states' assessments.

would need a degree of validity, but in the spirit of a general index would not have to be perfect, or near perfect, just generally representative of the nation's curriculum. The more general, the better in the sense that teaching to the test or focusing a school's curriculum on the test would be less desirable or practical.

The Achievement Index should be based upon age, not grade level. A national median for each age would be set in a baseline year, then used to set the index at 50. Then subsequent years would be reported as the percentage of students scoring above the average of the baseline year. For example, subsequent years could be reported as +2 or 52% - interpreted as 52% of the current students performed above the average level of the baseline year.

Using a percentage of students here and a percentage in the PACE Index has the advantage of referring to students as individuals within the educational system contrasted with comparing an average score that has relatively little intrinsic meaning. For example, a PACE index value of 83 would be interpreted as 83 out of 100 students are predicted to graduate—17 will not. An achievement index value of 57 means that 57 students out of 100 exceeded the baseline year's average.

This style of reporting, using a percentage of students, avoids some of the problems associated with percentiles, normal curve equivalents, grade equivalents, and standard scores, all of which appear to be somewhat abstract to parents and the general public, and in the opinion of educators, too open to misinterpretation.

These ideas are presented to kick off discussion. Why can't we have indexes for public education?

The general state of the art at the state level is very similar, except that there is less confidence in data quality across school systems and less sensitivity to the difficulty of adding data elements to existing collection processes.

The quest to identify all possible indicators for education has led to this conclusion: We have now made reporting the status of elementary and secondary public education so complicated that our audiences have gone from wanting more information to wanting to know which information is really meaningful to them. After years of hearing local school staff and members of the Board of trustees ask for more and more information, our local reporting must have caught up with and passed their ability to deal with everything, because within the past two years, the requests began to swing back toward asking for more concise summaries of the data. The ultimate resolution appears to be to have at hand the details or the ability to generate the details as needed, and to design better summaries that focus on the most frequently needed information. Re-read the Gardner quote on page 3.

ESP Insight

We have now made reporting the status of elementary and secondary public education so complicated that our audiences have gone from wanting more information to wanting to know which information is really meaningful to them.

Conclusion

Indicators and indexes can be developed and used to address each of the three hot issues identified at the beginning of this paper.

- Multiple Indicators
- Growth Models
- No Child Left Behind

Informing insight with indicators and indexes that are crafted to synthesize data for us is the goal. We can create a culture of data driven decision making in education by providing the right data, in the right way, at the right time. Over time, education will improve as an outcome of extraordinary insight.

Attachment A: Examples of Indicators and Indexes for Education

NOTE: The dissemination of an indicator as leading or trailing is dependent upon whether it is reported to forecast or to evaluate. A trailing indicator can become a leading indicator if it is reported in a timely manner with a useful periodicity.

Area: INSTRUCTION	Indicator / Index	Type of Indicator	Leading / Trailing
Graduate Preparedness Index (ACT/SAT, AP/IB, Advanced Course Completion, College Entrance Rate, State Exit Exam, College Remedial Course Enrollment)	Index	Outcome	Trailing
Opportunity for Quality Instruction Index (Student Average Daily Attendance, Teacher Attendance, Days of Normal Instruction)	Index	Process	Leading
AYP Index	Index	Outcome	Trailing
Proficiency on State Assessment	Indicator	Outcome	Trailing
Proficiency on Diagnostic Assessments	Indicator	Process	Leading
Completion of Remedial Courses	Indicator	Process	Leading
Tutorial Hours Funded	Indicator	Resources	Leading
Number of Students Enrolling Who Failed Prior State Assessments	Indicator	Context	Trailing

Area: FINANCE	Indicator / Index	Type of Indicator	Leading / Trailing
Financial Health Index (Audit Results, Fund Balance, Monthly Financials)	Index	Outcome	Trailing
Fund Balance	Indicator	Outcome	Trailing
Cycle Time for Accounts Payable	Indicator	Process	Leading
Tax Collection Rate	Indicator	Resources	Leading
Taxable Property Value	Indicator	Context	Leading

Area: LIBRARY/MEDIA SERVICES	Indicator / Index	Type of Indicator	Leading / Trailing
Collection Circulation Index (Books Borrowed, Media Borrowed, Time Items Kept, User Satisfaction, Increase in Circulation)	Index	Outcome	Trailing
Circulation	Indicator	Outcome	Trailing
Cycle Time to Purchase and Display New Books	Indicator	Process	Leading
Books and Media Displayed	Indicator	Resources	Leading
Age of Books and Media	Indicator	Context	Trailing

Area: FACILIITES	Indicator / Index	Type of Indicator	Leading / Trailing
Facility Usage Index (Percent Area in Use, Percent Area in Primary Use, Days Facilities Closed for Repair, Students per Square Foot))	Index	Outcome	Trailing
Days to Open New Facilities	Indicator	Outcome	Trailing
Cycle Time to Repair Facilities	Indicator	Process	Leading
Maintenance FTEs per Square Foot	Indicator	Resources	Leading
Age of Buildings	Indicator	Context	Trailing

Area: FOOD SERVICES	Indicator / Index	Type of Indicator	Leading / Trailing
Nutritional Value of Meals Index (Nutritional Content of Individual Meals, Type of Meals Served, Proportion of Each Type Served)	Index	Outcome	Trailing
Meals Served	Indicator	Outcome	Trailing
Cost per Meal	Indicator	Process	Leading
FTE Food Service Positions per Meal Served	Indicator	Resources	Leading
Number of National School Lunch Program Meals Eligibility	Indicator	Context	Leading

Area: HEALTH SERVICES	Indicator / Index	Type of Indicator	Leading / Trailing
Student Health Index (Immunizations, Absences, Referrals, Family Health Practices)	Index	Outcome	Trailing
Immunization Rate	Indicator	Outcome	Trailing
Office Visits by Students During Class Periods	Indicator	Process	Leading
Students per Nurse	Indicator	Resources	Leading
Percent of Students without Family Health Insurance	Indicator	Context	Trailing

Area: TRANSPORTATION	Indicator / Index	Type of Indicator	Leading / Trailing
Safety Index (Miles Driven, Accidents, Violations, Driver Experience, Driver Training)	Index	Outcome	Trailing
Miles without Accidents	Indicator	Outcome	Trailing
Driver Hours of Training	Indicator	Process	Leading
Substitute Driver Availability	Indicator	Resources	Leading
Fuel Efficiency of Existing Fleet	Indicator	Context	Leading

Area: SPECIAL EDUCATION	Indicator / Index	Type of Indicator	Leading / Trailing
IEP Success Index (IEP Objective Count, IEP Objective Completion Count, IEP Adjustment for Changes)	Index	Outcome	Trailing
Proficiency Rate on State Assessments	Indicator	Outcome	Trailing
Cycle Time from Referral to IEP Approval	Indicator	Process	Leading
Students with IEP per Speech Therapist FTE	Indicator	Resources	Leading
Number of Identified Students Projected to Enroll	Indicator	Context	Leading

Area: REMEDIAL PROGRAMS	Indicator / Index	Type of Indicator	Leading / Trailing
Students Promoted with Proficiency Index (Proficiency on State Assessments, Number of Retained or Promoted Students, Students Failing but Promoted by Committee)	Index	Outcome	Trailing
Students Meeting Individual Plan Objectives	Indicator	Outcome	Trailing
Cycle Time to Begin Services for Mobile Students	Indicator	Process	Leading
Tutorial Hours Available	Indicator	Resources	Leading
Number of English Language Learners New to School	Indicator	Context	Leading

Area: HUMAN RESOURCES	Indicator / Index	Type of Indicator	Leading / Trailing
Highly Qualified Teacher Index (Teacher Certifications, Core Courses Taught, Exceptions/Alternative Endorsements, Teacher Experience)	Index	Outcome	Trailing
Days Positions are Filled with Qualified Employees	Indicator	Outcome	Trailing
Cycle Time to Fill Open Positions	Indicator	Process	Leading
Automated vs. Manual Processes	Indicator	Resources	Leading
Teacher Turnover Rate	Indicator	Context	Trailing

National Forum on Education Statistics – Forum Guide to Education Indicators

The following pages extracted from the *Forum Guide to Education Indicators*, are used with permission from the National Forum on Education Statistics.

National Forum on Education Statistics. (2005). *Forum Guide to Education Indicators* (NFES 2005-802). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Chapter 1
INTRODUCTION TO EDUCATION INDICATORS AND INDICATOR SYSTEMS

This chapter defines the concept of an "education indicator" and describes the process of establishing a body of education performance and context indicators that will support decisionmaking by supplying useful, valid, reliable, timely, and cost-effective information.

What is an Education Indicator?

An education indicator is a measure of the status of, or change in, an educational system with regard to its goals. Examples include average student scores on assessments, graduation and completion rates, and teacher retention rates. Many users of education data focus on performance indicators—those indicators that measure the outcomes of the education system (i.e., student achievement and success). But because of the complexity of the education enterprise, sound education indicator systems must also include context indicators—those measures of system inputs and processes that aid in the interpretation of performance indicator (i.e., outcome) data. (See figure 1.)

An education performance indicator is a measure of the status of, or change in, an education system with regard to its goals.

The Role of Indicators in Education

Educators, parents, community members, politicians, business leaders, and the media use education indicators to compare schools against themselves over time, schools against peers (e.g., within a district or state), districts within states, and

Figure 1. Education indicator categories as system inputs, processes, and outcomes.

The advent of high stakes education indicators requires the generation of high quality indicator data. Thus, indicators should be

- > **useful** (i.e., relevant to the issues in question);
- > **valid** (i.e., measure what they purport to measure);
- > **reliable** (i.e., produce consistent measures over time);
- > **timely** (i.e., available in time to inform decisionmaking); and
- > **cost-effective** (i.e., produce information that is valuable enough to justify any collection burden).

states across the nation. To varying degrees, individual school leaders—including board members, superintendents, principals, and teachers—support the use of education indicators as the primary mechanism for measuring whether elementary and secondary schools are accomplishing stated goals. In fact, many local school boards, most state legislatures, and the federal government have established education accountability systems focused on achieving student and school performance targets. Those who favor using education indicators for accountability purposes argue that baselines, standards, and "hard" data are necessary to evaluate the status and progress of our education system and its "product," student learning.

While education leaders and policymakers appreciate the instructional and administrative need for the information provided by education indicators, not everyone has expertise in the development of useful, valid, reliable, and timely education indicators. Moreover, even properly constructed indicators may be misinterpreted. Evaluating education organizations based on sometimes confusing terms that are almost always inconsistently defined and interpreted may lead to unproductive comparisons of "apples to oranges." Rarely do such comparisons lead to the impartial, data-driven decisionmaking envisioned by educational and political leaders.

Selecting a body of education indicators should be driven by ongoing policy questions related to an education organization's goals and objectives.

Aligning Indicators with Policy Goals and Objectives

Although the use of indicators should be driven by policy needs, an indicator system does not need to answer every policy question. In fact, the considerable effort required to develop and refine indicators is warranted only to address ongoing policy needs rather than to answer infrequent or even one-time questions.

Aligning Indicators with Policies, Goals, and Objectives

Policy Goal: Prepare students to lead successful lives.

Assumption: To lead successful lives, students need to earn a high school diploma.

Objective: Focus support on keeping students in school.

Performance Indicators needed to assess efforts:

- Performance Indicator = *High School Graduation Rate*
- Performance Indicator = *High School Completion Rate*
- Performance Indicator = *High School Dropout Rate*

Context Indicators that might inform the interpretation of performance indicators:

- Context Indicator = *Student Stability Rate*
- Context Indicator = *Student Truancy Rate*
- Context Indicator = *Student Course Enrollment*

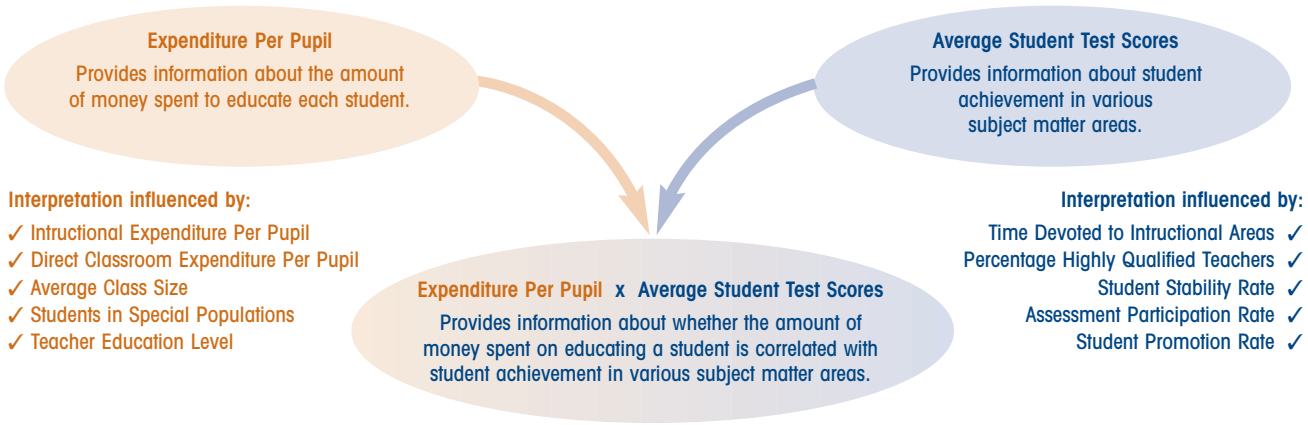
Bodies of Indicators

Not all indicators provide information about educational performance. Therefore, input and process indicators—context indicators—are sometimes needed as building blocks for performance indicators. They also offer insight into the interpretation of performance indicators. Because a single education indicator cannot possibly supply all the information needed to assess the status of, or change in, an education system, painting an accurate picture of the K–12 education enterprise requires a body of both performance and context indicators, with each individual indicator imparting a complementary piece of the puzzle. When a body of indicators is developed, the result is a well-integrated, multidimensional indicator system whose total value is greater than the sum of its parts.

For example, figure 2 illustrates how information provided by context indicators would enhance the interpretation of two commonly used performance indicators, Expenditure Per Pupil and Average Student Test Scores. Taken individually, or even in combination, Expenditure Per Pupil and Average Student Test Scores provide a limited perspective on what is happening in an education system. A more complete and accurate understanding requires additional information provided by context indicators.

Suppose two schools had roughly the same Average Student Test Scores but radically different Expenditures Per Pupil. In the absence of any additional context, you might infer that the school with the lower Expenditure Per Pupil was run more efficiently. From a policy perspective, you might even determine that this school should serve as a model for the one that spent more money per student to achieve the same results. While this might be true, student demographics could explain the difference as well. For example, a small student population will increase a school's Expenditure Per Pupil because there are fewer students to absorb overhead costs (for example, a school needs to have a principal regardless of how many students it has). Another explanation for the discrepancy between the two schools could be found by comparing indicators describing differences in teacher characteristics (for example, teachers with more experience usually earn more than teachers with less experience). Or maybe resource supplements from a parent–teacher organization could account for some of the disparity. In other words, the environments in which schools function may vary substantially. Therefore, any information that adds context or meaning to the performance indicators will lead to more appropriate data interpretation.

Figure 2. Context indicators.



Graphical presentation of how the combination of multiple context indicators is necessary to interpret the meaning of even apparently straightforward performance indicator findings.

Numbers versus Information

A high Total Expenditure Per Pupil may be interpreted as wasteful spending. Another explanation could be a large population of special needs students. Or, perhaps, a small total student population has required basic overhead costs associated with running a school to be distributed over a smaller number of students. Either way, these circumstances may greatly affect the interpretation of this commonly used indicator. They also illustrate the importance of using additional context measures to interpret education indicator data.

Using Context to Interpret Indicator Values

Indicators are value neutral until interpreted in light of their context. For example, "air temperature" is an indicator that we use every day, but it is just a number unless put into context. After all, 50 degrees would be considered quite balmy on a January day in Boston, whereas the same 50 degrees in Los Angeles in June would be considered unseasonably cold. Value judgments about indicators (e.g., whether it is warm or cold when it is 50 degrees) are external to indicator measurements and generally are assigned during interpretation rather than during collection.

School and community leaders should therefore consider organizational context when establishing appropriate goals and targets for indicator values. Doing so demands a thorough understanding of both indicators and the organization. For example, what is the "ideal" value for Teacher Class Absence Rate? Zero percent might be an immediate, but ill-considered, reply. After all, teachers get sick like everyone else. Moreover, teachers need professional development to improve their skills, and this will also cause them to miss class on occasion. Indicator developers and interpreters must somehow establish goals that take into account the real-world context in which their data are generated. Until they do, an indicator is only a number whose significance may or may not be properly interpreted.

> *The greatest danger of indicators is the ease with which they can give false impressions because they are misunderstood or interpreted in invalid ways.*
>
> —*Accountability Mechanisms in Big City School Systems* (ERIC/CUE Digest No. 71.).

Context indicators can also provide a system of checks and balances within an accountability reporting effort. For example, a school might be able to report improved achievement results if its low-achieving students do not participate in an assessment (perhaps because they were retained in a class that was not tested). This practice might go undetected unless nonachievement data, such as class enrollment, are available to provide additional context about school operations and processes. This phenomenon of improving one indicator value at the expense of another (intentionally or not) can be quite real and may certainly affect the interpretation of indicator meaning.

Finally, planners must also be aware that an organization's context may change over time. Student and staff characteristics change under many circumstances, and this is especially true in organizations with high student mobility and staff turnover rates. Curriculum may also change. Community resources may increase or decrease as overall economic conditions fluctuate. In addition to actual changes to traditional context indicators, expectations can also change. Meeting this year's target might be interpreted as a victory, but hitting the same target three years in a row could be viewed as stagnation.

Because context is so important in the interpretation of education indicators, this document includes context indicators as a vital component of any comprehensive education indicator system (figure 3).

Although this document focuses on indicators derived from administrative records systems, other resources may provide valuable data as well. For example, the *Youth Risk Behavior Surveillance System* survey provides information to policymakers that cannot otherwise be gathered by administrative records systems.

The Balloon Effect

Push a system in one place and it will expand in another place. This "balloon" effect can occur when an organization improves one indicator value at the expense of another, whether intentionally or not. For example, a school that successfully implements a policy to keep students from dropping out may, in fact, decrease its dropout rate (a good thing); but it should not be surprised to find a related decrease in the percentage of students going to college (not such a good thing, without an explanation).

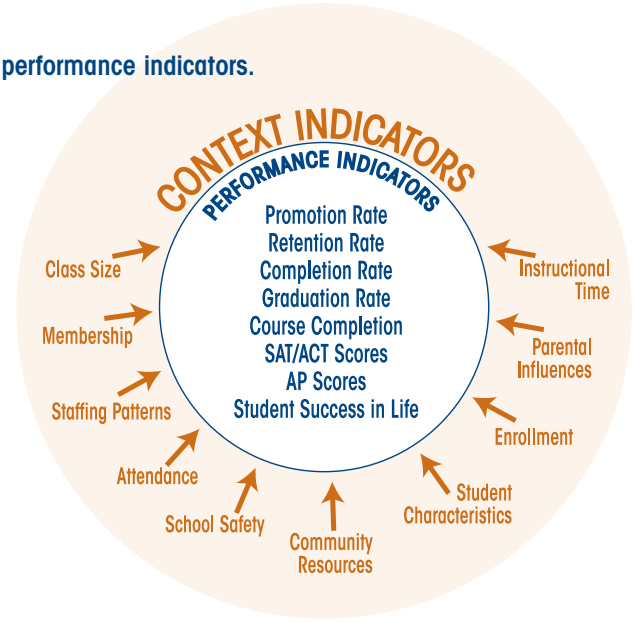
Generating Indicators

In our education data system, information is collected, analyzed, and reported at the local, state, and national levels, most frequently via transfer from the schools and school districts where it is collected to state education agencies and the federal government. Each level of administration has a different need for the information. Schools and school districts certainly need longitudinal records for individually identifiable students to monitor and evaluate the educational services they provide. In recent years, however, many state education agencies have also moved to a data model that benefits from maintaining individual unit records for students and staff. Unit records provide high data resolution when needed (e.g., for tracking highly mobile students between districts and verifying data submissions), while also allowing for aggregation when analysis and reporting do not require (or permit) individual identification. In fact, the vast majority of public reporting by state agencies occurs at the school and district level in the form of aggregate student information (e.g., the number of students completing high school in a given school or district). This emphasis on aggregate data is virtually complete by the time education data reach the U.S. Department of Education, the Bureau of the Census, and other federal agencies. Nonetheless, nationally reported aggregates are based on data that originated in individual student records in schools and classrooms across the country.

In general terms, a "system"—a national education data system as described above, an ecosystem, or even our solar system—is merely a set of regularly interacting parts that form a unified whole. Within an administrative records system in

Figure 3. Context versus performance indicators.

Examples of context indicators that might affect the interpretation of performance indicators.



An aggregate statistic in one data system may be a data element in another. For example, a count of student membership in a school district is an aggregate of individual school counts that reside in the district data system; this same number stands alone as a data element in a state education agency data system.

Similarly, an aggregate may become a data element even within the same data system. For example, Average Daily Membership in a school district might be calculated once for an academic year, then stored as a data element for future use.

an education organization, the "parts" have the highest resolution at the level of the "data element." A data element is the lowest level of information (i.e., data) that gets stored. For example, within a student record system (see figure 4), Quiz 1 Student Score is (1) a singular data element for an individual student on a given quiz. This element may then be (2) aggregated for multiple students, creating Quiz 1 Class Average Score, (3) calculated over time, creating Average Score on Quizzes for a Student, (4) combined to derive Class Average Score and, ultimately, (5) configured to create the performance indicator Average Student Score.

In this way, a piece of data in an individual student, staff, or education institution record may be aggregated, combined, and calculated until new information is derived. When managed in a consistent fashion, each of these pieces of new information may, in turn, be scaled by class, school, district, state, and country to add other dimensions of analysis from a single data element. The key to aggregating data from different records is collecting the data in a consistent manner. While it might be reasonable to assume that the registrar, teacher, or school secretary who records student attendance in a school does so consistently from day to day, month to month, and year to year, the same is not necessarily true across schools, districts, and states throughout the nation. For example, some education institutions might report that individual students attended school on a given day only when they were present for four or more hours. Other institutions might define "attendance" as a student being present at any point during the school day. Clearly, the significance of a "95 percent average daily attendance rate" would then vary substantially, which illustrates the need for standard data element and indicator definitions if statistics from different organizations are to be compared.

The *Forum Guide to Building a Culture of Quality Data: A School and District Resource* provides best practice recommendations for generating high quality data in schools. It is available at no cost at *http://nces.ed.gov/forum/pub_2005801.asp.*

Ideal Indicators and Real World Tradeoffs

The quality of an indicator cannot surpass the quality of its components (data elements). In addition to normal data quality issues (e.g., student misreporting, entry errors, and changing assessment tools), at least two other factors affect the quality of education indicators: (1) the complexity of the issue being measured and (2) previous experience studying the issue *(Monitoring School Quality: An Indicators Report,* NCES 2001). One way of assessing the quality of an education indicator is to consider its utility, validity, reliability, timeliness, and cost-effectiveness:

Utility

An indicator should be useful for answering, or helping to answer (as with a context indicator), an important policy question. If the indicator does not provide

Figure 4. Turning data elements into indicators.

QUIZ SCORES (in percent correct)

	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Average Score
Jose	95 [1]	95	90	95	90	93
Emma	95	90	75	85	85	86
Angelique	90	100	85	90	90	91
Amy	95	80	95	75	80	85
Terrence	95	95	90	90	95	93 [3]
Nancy	90	90	90	95	100	93
Paul	100	75	90	90	85	88
Audrey	100	90	90	85	90	91
Omar	90	85	75	80	95	85
Maria	100	100	85	90	90	93
Class Average	95 [2]	90	86.5	87.5	90	89.8 [4] = Class Average Score
						89.8 [5] = Average Student Score

1> Singular data element
2> Aggregated (column calculation)
3> Calculated (across columns)
4> Derived (rows and columns)
5> Performance indicator

A single data element (Quiz 1 Student Score[1]) may be aggregated (Quiz 1 Class Average Score[2]), calculated over time (Average Score on Quizzes for a Student[3]), combined to derive new information (Class Average Score[4]), and formatted as a performance indicator (Average Student Score[5]).

useful information, there is little reason to collect or calculate it, even if doing so can be accomplished accurately, reliably, cost-effectively, and in a timely manner.

Validity

> *"To say that any important educational outcome is measurable is not to say that satisfactory methods of measurement now exist."*
>
> —**Robert Ebel, *Practical Problems in Educational Measurement***

Validity is the degree of correspondence between a measurement and the process or product being studied. In other words, is the indicator accurate? Does it measure what it purports to measure? Is it free from bias (i.e., a systematic error in data generation or collection)? If so, the indicator is considered valid.

Reliability

Reliability refers to a measure's consistency, reproducibility, and dependability. In other words, if the same indicator were to be measured multiple times, would the same results be generated? Without consistent measurement methods, results from

How many paper towels per student mile bused does your district use?

Most state education agencies (SEAs) publish vast amounts of data about their schools and districts. A taxpayer advocacy organization in one state cross-tabulated all the data published by its SEA in order to create a volume of indicators for measuring and comparing the state's schools and districts. Because the organization compared all available data elements and aggregates, the report included meaningless "indicators" such as Paper Towel Use Per Student Mile Bused.

Education indicators grow out of data elements (figure 4), the lowest level of data stored in an administrative records system. But not every data element may be aggregated, calculated, or combined into a piece of information that helps an education organization measure its status or progress in a meaningful way.

Haphazard approaches to producing indicators confuse student and school assessment rather than illuminate it, illustrating that considerable thought must go into developing useful education indicators.

Just because technology enables exhaustive comparisons of data in administrative records systems does not mean that doing so in a random manner is helpful. In fact, it may cloud rather than clarify our understanding of how our schools and students are performing.

> **Straightforward indicators such as Average Teacher Experience and Average Class Size usually generate high quality data, as do measures that have been studied for a long time, such as Assessment Score Results. Data about new areas of interest, however—including professional development, student discipline, and technology availability—generally produce lower quality data, as do particularly complex topics such as "leadership" and "pedagogy."**
>
> —*Monitoring School Quality: An Indicators Report* (NCES 2001)

different organizations or even from within the same organization at different points in time (e.g., longitudinal or time-series data) cannot be compared. Standard or "best practice" collection methods, therefore, are vital to any data/indicator system from which information will be drawn for the purpose of making comparisons (e.g., among groups of students, schools, school districts, states, pedagogical practices, reform strategies, or other entities). Similarly, comparing the progress of an individual or institution against itself over time is pointless without ensuring that the measurement practice itself has not changed (i.e., it is reliable).

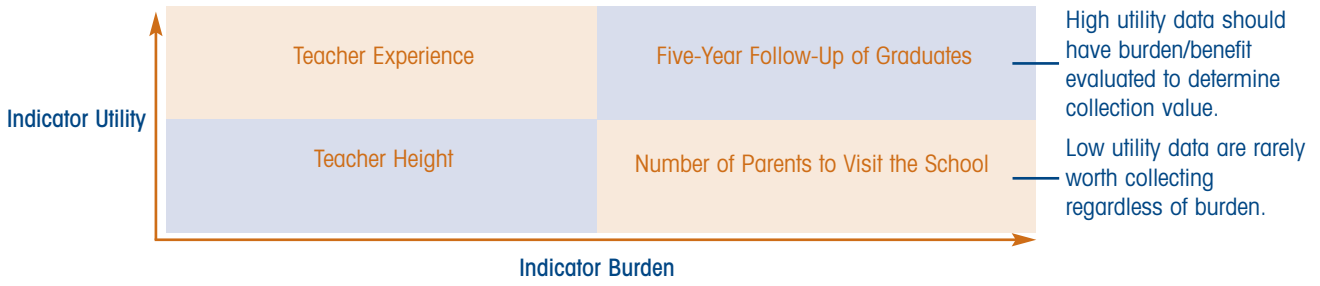
Timeliness

Data are most valuable when they are readily available for informing decisionmaking, which means that the data have to arrive in time to influence decisionmakers (e.g., board members should receive fiscal data in time to inform budget planning just as teachers should receive test results in time to inform instructional planning). At the same time, policymakers should not be hasty in using indicators to make changes. Change should not be haphazard. Trends in indicator data are far more revealing than findings for a single year. Thus, changes to policies and procedures should be in response to indicator trend analysis rather than inferences about one or two years' worth of indicator results.

Cost-Effectiveness

Although what is easy to collect may not be what is desirable to collect, expensive data collections such as one-time surveys are usually not appropriate as inputs into performance indicators that will be generated year after year. Conversely, some good data (i.e., important or necessary to have) that are expensive to collect are still worth the effort (see figure 5). Data "burden" may be defined as collecting or manipulating data for a requester in a way that is of little value to, or demands significant financial or human resources from, the provider.

Figure 5. Indicator data collection: utility versus burden.



Indicator utility should be weighed against burden to determine the relative benefit of collection.

Data burden may arise in at least two dimensions:

Attitudinal: "Why do they want me to collect and report it this way when doing so doesn't help me and I am so busy?"

Financial: Money spent directly on a collection or indirectly on staff time and equipment to support a collection.

In our computer-driven world, even the transfer of electronic records to paper may be perceived as a burden.

This burden to the provider may be measured in terms of both time and real dollars spent on the endeavor. For example, burden arises when a data provider is asked (or required) to complete multiple data requests, each for the same type of information in a slightly different form—after all, the utility of the information is not improved for the provider by taking time to modify its format solely to meet the needs of an external request. Placing a burden on data providers should be of great concern to data requesters. Data providers who see little value in a collection may be less likely to allocate the resources necessary to ensure they are collecting and reporting high quality data.

Making Tradeoffs

Ideally, indicators in an education data system are useful, valid, reliable, timely, and cost-effective. But these, too, are relative terms. What is useful for one organization may not be so for another. Similarly, different organizations might have differing definitions of what is cost-effective or burdensome. While sharing information and expertise is always a good idea, an education institution, with its own goals, priorities, policies, and circumstances, simply cannot borrow a list of indicators from its neighbor, even if the institutions are peers in many respects. Instead, data- and policy-leaders must consider the information needs unique to the organization's goals and priorities and, subsequently, the costs and benefits of indicators that may meet those needs. For example, if an indicator is particularly useful (e.g., Five-Year Follow-Up of Graduates), maybe a relatively substantial burden is worthwhile. Perhaps it is not, however, if the indicator's validity or reliability is questionable, or if data needed to generate the indicator are not available in time to be useful. When real world constraints kick in, as they generally do, tradeoffs between quality, reliability, and utility become inevitable. The job of the indicator system development team is to ensure that these tradeoffs are made reasonably and responsibly (see figure 6).

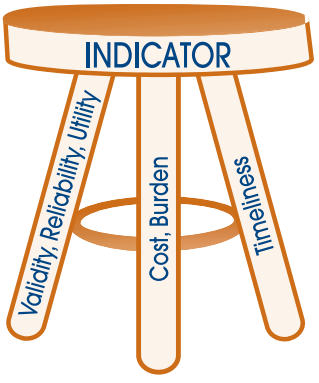


Figure 6. Selecting good indicators. Selecting "good" indicators is like building a stool. Consideration must be given to the three legs: Quality (Validity, Reliability and Utility), Burden, and Timeliness. In the real world, tradeoffs may be necessary, but at least two of the characteristics must always retain their integrity, with as little compromise to the third as possible to keep the indicator feasible.

Opening a dialogue between those who make policy and those who develop indicators should:

- inform indicator developers about policy goals, objectives, and evaluation strategies; and
- inform policymakers about the capabilities and limitations (including a cost-benefit analysis) of various indicator options.

Who Develops Indicators

Education indicators should be developed by people who understand the:

- institution's policies, goals, and objectives;
- information needed to evaluate the status of, or progress toward, those goals and objectives;
- capabilities and limitations of the organization's data system;
- external reporting demands (e.g., to the school district, state, or federal government); and
- best practices for selecting and developing education indicators.

Unfortunately, very few individuals have mastered all these areas of expertise. On the positive side, however, staff who possess a thorough understanding of the organization's data system, its external reporting responsibilities, and best practices for selecting and developing education indicators may be available. Their job becomes to communicate with the policymaking personnel in the organization to ensure that they (the indicator developers) learn about policies, goals, and objectives and their corresponding information needs. In other words, indicator producers and indicator users must engage in a discussion. Politicians, policy analysts, board members, senior administrators, and researchers must explain their information needs to those responsible for developing the indicator system. Conversely, as the indicator developers improve their understanding of the policymakers' information needs, they should, in turn, explain the costs, benefits, implications, and limitations of alternative approaches to producing indicators (see Making Tradeoffs).

For example, suppose school leaders set as a priority the improvement of student academic performance on state assessments, and that one approach to attaining this goal would be to decrease class sizes in an effort to enrich student–teacher interaction. Once the indicator development team understands this policy target (learned through dialogue with the policymakers), they may explore the policymakers' information needs: Are policymakers interested in Average Class Size, Maximum Class Size, or Median Class Size? Are they focused on core subject areas or all courses? Is the interest centered on elementary or high school classes?

Once these and other questions are answered, indicator developers will have a much better understanding of policy needs. They can then proceed to explain data options to the policymakers.

For example, they may suggest that class size be measured on a per-class basis, but as this is not an indicator the organization currently maintains, calculating it would carry both time and financial burdens. Policymakers might decide that assessing the initiative is worth the staff and monetary costs. Alternatively, they may choose to use Student:Teacher Ratio as a proxy for Average Class Size after the indicator developers explain that calculating it would be relatively burden-free (the organization already collects the data element components), it would provide a reasonable approximation of Average Class Size (average class size corresponds with the total-student-to-total-staff ratio, albeit not perfectly),

and it accurately reflects increases and decreases over time (even though it is only an approximation). If the indicator development team adequately explains the possible indicators and related data elements, policymakers should fully understand the utility, validity, reliability, timeliness, and burden aspects of their options. They should also understand the capabilities and limitations of the indicator data they eventually choose.

Unintended Consequences

Many people believe that indicators are simply numbers that represent something (e.g., a dropout rate) and that these numbers, on their own, do not inherently convey judgment or consequence. Rather, judgment and consequences are imposed when indicator values are interpreted and, in response, policies are changed. While this is true, the establishment of an indicator can nevertheless result in unintended consequences that can be very real. For example, indicators perceived to suggest inadequate student achievement might lead to a policy of high-stakes testing. While this may seem reasonable, this indicator-driven policy may have the unintended consequence of increasing student retention rates as teachers become more cautious about promoting students who may not fare well on high-stakes assessments. It could also lead to changes in course offerings (e.g., less time for subject matter not on the test) or other unintended outcomes.

While unintended consequences are very hard to predict, planners must nevertheless try to explore the potential ramifications of the indicators they use. Some ramifications may prove tolerable, others unacceptable. Either way, planners should proactively consider the desirable and undesirable effects of indicator use and policy response.

Other Important Best Practices

Training Users

As described above, dialogue between data and policy specialists is beneficial for an organization on several fronts. Indicator developers are better able to provide the right data for informing policy decisions when they learn about information needs directly from those responsible for making and evaluating policy. Moreover, policymakers learn something as well—the characteristics of the data, including capabilities and limitations—which minimizes mistakes in interpretation and use.

In addition to this initial dialogue, formal training for using and interpreting indicators is essential. Staff must be properly prepared to answer questions likely to arise about the data (e.g., when a parent or newspaper reporter calls). Indicator

Proxy Data

A "proxy" is basically a substitute for the real thing. For example, in education data, the element Free and Reduced Price Meal Eligibility is frequently used as a proxy for a student's status as economically disadvantaged. Admittedly, this (and every) proxy does not correlate perfectly with its principal information target; some children (especially high school students) choose not to participate in the meal program even though they are eligible. Nonetheless, Free and Reduced Price Meal Eligibility is a reasonable estimation of a student's status as economically disadvantaged—without carrying the burden of asking families to report their income.

Indicator systems that use proxy elements must confirm that these proxy elements relate to the principal issue they are meant to estimate. This does not mean that a school district must conduct original research to support the relationship, but the organization should be able to document the choice with relevant studies or other standards within the field. Moreover, the organization should be aware of any limitations associated with using the proxy data.

staff at the local and state education agency levels have an added training responsibility: they must ensure that data providers (e.g., school or LEA staff) are warned about the release of potentially controversial or otherwise high-profile public information. It often makes sense to embargo data for several days to a week (or even a month or longer for particularly high-stakes data) so that school and district staff may process the data's implications and prepare to deal with the sometimes very public ramifications of releasing indicator results (be they positive or not).

Reporting Indicator Data

Policies must be established to govern indicator reporting in order to protect the privacy of individual students and staff.

Indicator data must be reported and shared with the same cautions and concerns other education data warrant. If, for example, an education organization collects unit records (i.e., individual student or staff records) to generate an education indicator, the organization should still comply with the privacy guarantees afforded students and staff by local, state, and federal laws. Even if indicators are reported in aggregate form, commonly accepted statistical procedures must be followed to protect individuals from inadvertent identification. These standard procedures include cell size limitations to prevent the identification of individual students within small groups of "aggregates."

Furthermore, an organization should not tell its data providers it is collecting data for a specific purpose, then use the information for a different purpose. For example, an SEA should not tell LEAs it is collecting data for federal reporting, then use the data to compare LEA performance. In order to reduce the unexpected (or unacceptable) use of indicator results, any organization collecting data should develop and maintain policies governing the distribution and use of indicator data by its own staff as well as by outside agencies and organizations with access to the data.

The *Forum Guide to Protecting the Privacy of Student Information* provides best practice recommendations for ensuring the privacy of information collected for, and maintained in, student records. It is available at no cost at *http://nces.ed.gov/forum/pub_2004330.asp*.

Securing Data and Information

Information (data) is one of an education organization's most valuable assets.

Given the time, energy, and money that goes into collecting data, an organization's information system is one of its most valuable assets. Yet threats to an organization's data exist in the form of natural events (e.g., lightning strikes, floods, aging media), intentional acts of destruction (e.g., computer hacking, software viruses, dissatisfied employees), and unintentionally destructive acts (e.g., programming errors, spilled coffee). It is no understatement to suggest that the three fundamental goals of data security are especially applicable to high-profile and high-stakes indicator data. Organizations must ensure data:

- **confidentiality**—preventing unauthorized disclosure and use of information;
- **integrity**—preventing unauthorized creation, modification, or deletion of information; and
- **availability**—preventing unauthorized delay or denial of information.

The *Forum Unified Education Technology Suite* presents a comprehensive approach to acquiring, implementing, managing, securing and using technology in education settings. It is available at no cost at *http://nces.ed.gov/pubs2005/tech_suite/*.

Statistical Integrity and Public Presentation

Two other issues of great importance to developing and preparing indicator data are statistical integrity and public presentation. Because these issues are critical to the responsible management of indicator systems (but not directly within the scope of this *Guide),* they are addressed in detail in appendices to this document:

- **Appendix B: Statistical Terms and Concepts**
  Describes statistical terms and concepts commonly used to conceptualize, develop, and interpret education indicators.
- **Appendix C: Display and Presentation Options for Indicators**
  Offers guidance for preparing indicator reports and displaying indicator data.

Chapter 2
CATALOG OF EDUCATION INDICATORS

This chapter describes 44 education indicators commonly used to measure the status of, or change in, education institutions across the nation.

Detailed descriptions of education indicators comprise the bulk of this chapter. An alphabetical list of common education performance and context indicators is also provided, as is an index by topic area. As illustrated in figure 7 below, each indicator entry contains a definition, a recommended use, a policy question, caveats and cautions, additional information, related indicators, data

Figure 7. Illustration of the framework and conventions used to describe indicators.

Indicator Name Identifies the primary indicator name used in this document.

Cross References Lists any alternative name that is commonly used for the indicator, and is cross-referenced to the indicator in this *Guide*.

Definition Describes or defines the indicator.

Recommended Uses Provides suggestions for the appropriate application of the indicator.

Policy Questions Identifies one or more broad policy questions that may be informed by the indicator.

Caveats and Caution Introduces specific issues to be considered in order to avoid misapplication or misinterpretation of the indicator.

Additional Information Explains other issues that may influence the use of the indicator.

Related Indicators Identifies other indicators that may be used in conjunction with the indicator to provide additional information.

Components Lists the data elements and aggregate, calculated, or derived statistics needed to generate the indicator.

Formula Provides the actual mathematical formula for computing the indicator value.

Commonly Reported Subgroups Identifies subgroups commonly reported for the indicator.

Display Suggestions Presents recommendations about the types of graphical or tabular display that are appropriate for the indicator (see appendix C).

Sample indicator layout

Assessment, Percentage Student Participation

Participation Rate in Student Assessment
Percentage Students Participating in State Assessment
Student Participation Rate in Assessment

Definition
A measure of student participation on an assessment (i.e., the percentage of students taking a test). "Participation" is often measured by the number of exams generating a valid score.

Recommended Uses
This indicator may be used to identify whether all students participated in an assessment. It may also be used to determine whether performance results might be biased (e.g., if an unusually high number of students did not take the exam).

Policy Question
Are assessment results based on a fair picture of students in our school or district?

Caveats and Cautions
> This indicator does not distinguish between students who were tested without modifications and students who took alternative assessments or for whom special test-taking modifications were allowed.
> Not all states use equivalent definitions to identify students eligible for participation in state assessments.
> Under some circumstances, organizations may choose to report, in the denominator, the number of "eligible" test takers (e.g., excluding students enrolled in the school or district less than a full academic year, or non-English speaking students in the United States for one year or less).
> In some jurisdictions, parents may refuse to allow their children to participate in assessments.

Additional Information
> This indicator does not distinguish between students who were tested without modifications and students who took alternative assessments or for whom special test-taking modifications were allowed.

Related Indicators
Adequate Yearly Progress (AYP), Percentage Schools Making
Adequate Yearly Progress (AYP), Percentage Schools in Improvement Categories
Adequate Yearly Progress (AYP), Percentage Local Education Agencies (LEAs) in Improvement Categories
Assessment, Average Student Score
Assessment, Percentage Students Demonstrating Proficient or Advanced Performance

Components *Italicized terms are defined in chapter 2, appendix A or appendix D*
Components include the total number of assessments generating a valid score (i.e., the total number of test takers) and the total enrollment on the date of test administration.
Numerator: Number of assessments generating a valid score
Denominator: Enrollment on the date of testing.

Formula
Assessment, Percentage Student Participation is calculated by dividing the number of assessments generating a valid score (i.e., the total numbers of test takers) by the total enrollment on the date the test was administered, and multiplying by 100 to create a percentage value.

$$\frac{\text{Number of assessments generating a valid score}}{\text{Student enrollment on test date}} \times 100$$

Commonly Reported Subgroups
Age, grade level, disability status, economic disadvantage status, English proficiency, migrant status, race, sex, and full- versus part-academic year enrollment status.

Display Suggestions
Generally presented as a number in the form of XX.X percent, this indicator may be displayed in tables or bar charts by subgroup. Each subject matter area is usually shown separately, but rates from multiple subjects may be compared in the same table or graph.

element components, a formula, commonly reported subgroups, and display suggestions. In addition to an alphabetical listing, the 44 indicators in this *Guide* have been indexed based on the following major policy and content strands:

Inputs
- ✓ Student/School Characteristics
- ✓ Financial Resources
- ✓ Staff Characteristics

Processes
- ✓ School Climate
- ✓ Opportunity to Learn

Outcomes
- ✓ School Performance

Table 1. Index of indicators by policy strand.

Page	Indicator Name (Alphabetical)	Common Name	Inputs			Processes		Outcomes
			Student/School Characteristics	Financial Resources	Staff Characteristics	School Climate	Opportunity to Learn	School Performance
19	Absence Rate (Class), Teacher	Class Absence Rate (Teacher)			✓	✓	✓	
21	Adequate Yearly Progress (AYP), Percentage Local Education Agencies (LEAs) in Improvement Categories	Percentage LEAs in Adequate Yearly Progress (AYP) Improvement Categories						✓
23	Adequate Yearly Progress (AYP), Percentage Schools in Improvement Categories	Percentage Schools in Adequate Yearly Progress (AYP) Improvement Categories						✓
25	Adequate Yearly Progress (AYP), Percentage Schools Making	Percentage Schools Making Adequate Yearly Progress (AYP)						✓
27	Alcohol-Related Incidents Reported Per 100 Students	Reported Alcohol-Related Incidents Per 100 Students				✓		
29	Assessment, Average Student Score	Average Student Assessment Score						✓
30	Assessment, Percentage Student Participation	Percentage Students Participating in State Assessment						✓
31	Assessment, Percentage Students Demonstrating Proficient or Advanced Performance	Percentage Students Demonstrating Proficient or Advanced Performance						✓
33	Certification, Percentage Classes Taught by Teachers Holding Emergency, Provisional, or Out-of-Field	Percentage Classes Taught by Teachers Holding Emergency, Provisional, or Out-of-Field Certificates			✓		✓	
35	Class Size, Average	Average Class Size		✓		✓	✓	
37	College Entrance Testing, Percentage Graduate Participation	Percentage Graduate Participation in College Entrance Testing						✓

Page	Indicator Name (Alphabetical)	Common Name	Inputs			Processes		Outcomes
			Student/School Characteristics	Financial Resources	Staff Characteristics	School Climate	Opportunity to Learn	School Performance
39	College Entrance Testing, Student Average Score	Average Student Score on College Entrance Testing						✓
41	Courses (Advanced), Percentage Student Completion	Percentage Students Completing (Advanced) Courses					✓	✓
43	Courses (Advanced), Percentage Student Enrollment	Percentage Students Enrolling in (Advanced) Courses					✓	✓
45	Criminal Offense Incidents Reported Per 100 Students	Reported Criminal Offense Incidents Per 100 Students				✓		
47	Drug-Related Incidents Reported Per 100 Students	Reported Drug-Related Incidents Per 100 Students				✓		
49	Education Level, Mother	Mother's Education Level	✓					
51	Education Level, Teacher	Teacher's Education Level			✓			
53	Experience Level, Teacher	Teacher's Experience Level			✓	✓	✓	
55	Expulsion Incidents Per 100 Students	Number of Expulsion Incidents Per 100 Students				✓	✓	✓
57	High School Completion/ Graduation Rate, Cohort Rate	Completion Rate/ Graduation Rate	✓				✓	✓
59	High School Completion/ Graduation Rate, Leaver Rate	Completion Rate/ Graduation Rate	✓				✓	✓
61	High School Dropout Rate, Annual Student	Annual High School Dropout Rate	✓				✓	✓
63	High School Dropout Rate, Cohort Rate	Cohort High School Dropout Rate	✓				✓	✓
65	"Highly Qualified" Teachers, Percentage	Percentage "Highly Qualified" Teachers			✓		✓	
67	Instructional Time, Allotted	Allotted Instructional Time				✓	✓	
68	"Persistently Dangerous" Schools, Percentage	Percentage Schools Identified as "Persistently Dangerous"				✓		
69	Placement of Students With Disabilities	Disabilities, Placement of Students With	✓			✓	✓	
72	Promotion Rate, Student	Retention Rate (Student)						✓
73	"Qualified" Paraprofessionals, Percentage	Percentage "Qualified" Paraprofessionals			✓		✓	
75	Retention Rate, Teacher	Teacher Retention Rate			✓			
77	School Capacity, Percentage Used	Percentage School Capacity Used	✓	✓				

Page	Indicator Name (Alphabetical)	Common Name	Inputs			Processes		Outcomes
			Student/School Characteristics	Financial Resources	Staff Characteristics	School Climate	Opportunity to Learn	School Performance
79	Stability Rate, Student Enrollment	Student Enrollment Stability Rate	✓			✓	✓	
80	Student:Instructional Computer Ratio	Instructional Computer:Student Ratio		✓			✓	
82	Student:Staff Ratio	Student:Teacher Ratio		✓		✓	✓	
84	Suspensions (Out-of-School), Actions Per 100 Students	Number of Out-of-School Suspension Actions Per 100 Students				✓	✓	✓
86	Suspensions (Out-of-School), Average Duration	Average Duration of Out-of-School Suspension Incidents				✓	✓	✓
88	Suspensions (Out-of-School), Percentage Students Receiving	Percentage of Students Receiving Out-of-School Suspensions				✓	✓	✓
90	Teacher:Administrator Ratio	Administrator:Teacher Ratio		✓				
92	Transportation Services, Percentage Students Receiving	Percentage Students Receiving Transportation Services	✓	✓				
93	Truancy Rate, Schoolwide	Schoolwide Truancy Rate				✓	✓	✓
94	Violent Incidents Reported Per 100 Students	Reported Violent Incidents Per 100 Students				✓		
96	Vocational/Technical Programs, Percentage Non-traditional Completers	Percentage Nontraditional Completers of Vocational/ Technical Programs					✓	✓
98	Vocational/Technical Programs, Percentage Non-traditional Participation	Percentage Nontraditional Participation in Vocational/ Technical Programs					✓	

ESP Optimal Reference Guides and Optimal Reference Books

ESP covers a wide variety of education topics with our series of informational whitepapers called Optimal Reference Guides (ORGs) and Optimal Reference Books (ORBs). All are available for free download at *www.espsolutionsgroup.com/resources.php*. You can also subscribe to our monthly newsletter to have ORGs and ORBs emailed to you as soon as they are published. Just visit the link above for more information.

Data Quality
- The Data Quality Imperative, Data Quality Series—Part I
- The Data Quality Manual, Data Quality Series—Part II

Data Management
- Actions Speak Louder than Data
- From Information to Insight—The Point of Indicators
- Aligning Indicators and Actions
- Data Management Strategy for States and Districts
- Defining Data
- Management of a Education Information System
- Our Vision for D3M
- Using Assessment Results to Get Performance Results
- Why Eva Baker Doesn't Seem to Understand Accountability—The Politimetrics of Accountability

Longitudinal Data Systems
- D3M Framework for Building a Longitudinal Data System
- The Dash between PK and 20: A Roadmap for PK-20 Longitudinal Data Systems
- What's Really "In Store" for Your Data Warehouse? Data Warehouse Series—Part I
- What's Behind Your Data Warehouse, Data Warehouse Series—Part II
- Accessing Student Records in a State Longitudinal Database, Data Warehouse Series—Part III

Project Management
- Why 70% of Government IT Projects Fail, Project Management Series—Part I
- From Risk to Reward: A Guide to Risk Management, Project Management Series—Part II
- Marketing Your Field of Dreams, Project Management Series—Part III

Electronic Transcripts
- Electronic Student Records and Transcripts: The SEA Imperative
- Why Your State Needs a PK-20 Electronic Record/Transcript System

Standards
- Articulating the Case for Course Numbers
- Confidentiality and Reliability Rules for Reporting Education Data
- FERPA: Catch 1 through 22
- Graduation Rates: Failing Schools or Failing Formulas?
- National Education Data Standardization Efforts
- Racial/Ethnic Data Reporting in Education
- Recommended Data Elements for EDEN Reporting
- Revisions to FERPA Guidance

Trends in Education
- Data-Driven Decision Making 2016
- How Education Information Fared in the Last Decade
- IT Defined…for the Educator
- Why My Space Matters to the K-12 Space

Student/Staff Identifiers
- Requirements for an RFP for Student Identifiers
- Statewide Student Identifier Systems

Disaster Prevention & Recovery
- Disaster Prevention and Recovery for School System Technology

Growth Models
- Growth Model Growing Pains, Growth Model Series—Part I
- Comparison of Growth and Value-Add Models, Growth Model Series—Part II
- Making a Year's Growth and Performing on Grade Level: Muddled Definitions and Expectations, Growth Model Series—Part III
- Growth Models—Finding Real Gains

ESP Solutions Group

(512) 879-5300

www.espsolutionsgroup.com