

The Optimal Reference Guide:
**From Information to Insight
—the point of indicators**

Extraordinary insight™ into today's education information topics

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

With a foreword by C. Jackson Grayson, Ph.D.



ESP Solutions Group

Table of Contents

Foreword.....	3
What’s the Point of Indicators?	9
Action Reports	9
Benchmarking vs. Thresholds	12
Indicators and Indexes.....	12
Index = Synthesis of Data (Multiple Indicators)	13
Formulating an Index	14
Converting Trailing Indicators to Leading Indicators	17
Vital Signs	17
What are the Indicators?	20
No Child Left Behind — AYP Index	20
Multiple Indicators.....	23
Accountability vs. Diagnostics	24
Growth Models.....	26
An Academic Growth Index	27
Not a Fan of Newsweek’s Rankings	27
Indicator vs. Public Index.....	27
Taxonomy of Indicators.....	29
Possible Public Indexes for Education.....	30
Pace toward Graduation	30
Student Mobility	31
Achievement Index.....	32
ATTACHMENT A.....	34
Examples of Indicators and Indexes for Education.....	34



Foreword

C. Jackson Grayson, Jr.
Chairman and CEO, APQC

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 it 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, and 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!

From Information to Insight — the point of indicators

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.



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.



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.

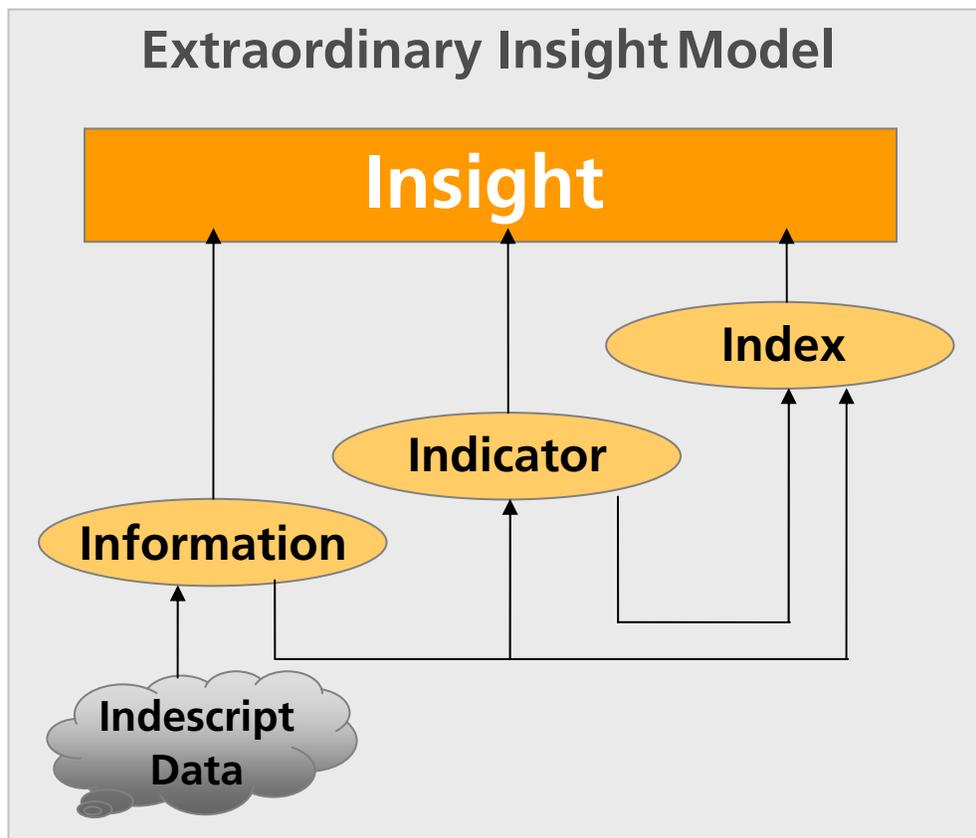


Figure 1: Extraordinary Insight Model

Indescript 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).

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



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

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.

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.

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.

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.

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

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

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.

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.

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



ESP Insight

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



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 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:

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.

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.

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

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

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.

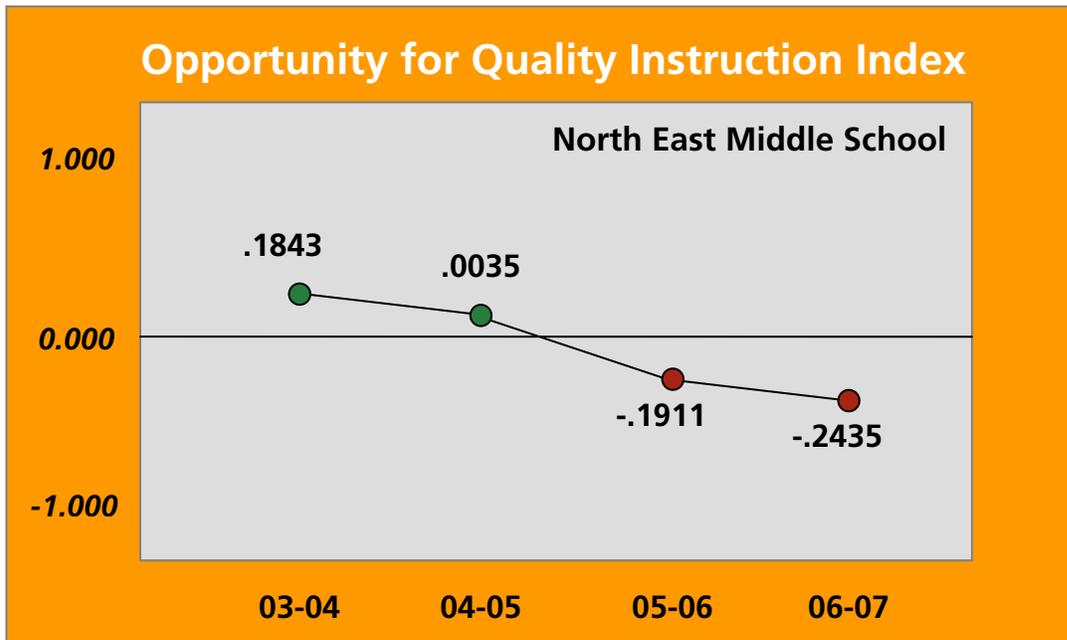


Figure 3: Opportunity for Quality Instruction Index

School: North East Middle School	Index: Opportunity for Quality Instruction	Current Index Value: 06-07: -.2435	Trend: 03-04: .1843 04-05: .0035 05-06: -.1911	Comment: This school's index value has declined over the past four years from being above average to being below average.
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, 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.

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?



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.



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.

- 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 PIIIE 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.

Two lessons learned from this experience are:

1. Simplicity communicates.
2. 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.)

 **ESP Insight**
Simplicity communicates.

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

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 be 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.*

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.



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.

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?

 **ESP Insight**
Graduation from high school is the quintessential indicator.

- 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.
- 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.

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

 **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.

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.

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

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.

 **ESP Insight**
Legislatures want accountability. Teachers want diagnostics.

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.

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

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.

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

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.



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	<i>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.</i>				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

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

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

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 public 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.

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.

 **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.

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.

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

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

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

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

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 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.



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: Facilities	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



About ESP Solutions Group

ESP Solutions Group provides its clients with *Extraordinary Insight*™ into K-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)**, **Education Data Exchange Network (EDEN)**, 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, contact Greg Nadeau at (512) 879-5346 or gnadeau@espsg.com.

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

From Information to Insight—the point of indicators, Copyright © 2007 by ESP Solutions Group. All rights reserved. No part of this paper shall be reproduced, stored in a retrieval system, or transmitted by any means, electronic, mechanical, photocopying, recording, or otherwise, without written permission from the publisher.



ESP Solutions Group

(512) 879-5300

www.espsolutionsgroup.com

Austin • Boston • Washington DC