

# Clear Answers to Common-Sense Questions about Accountability

by Dr. Douglas Reeves  
Thrust for Educational Leadership (March/April 2000)

*Parents and policy-makers have a direct agenda when it comes to accountability. Their questions must be addressed with clarity and candor.*

George Orwell (1949) helped generations of students learn the meaning of the term "double-speak." Few practitioners of the art of double-speak have studied the concepts as assiduously or applied it with as much vigor as attackers and defenders of public education. Both the strident tone of the debate between these two camps and the obscurity of the arguments employed by the debaters cause many observers to throw up their hands in dismay.

Parents and policy-makers have a more direct agenda. In terms of educational accountability, they express the issues very clearly.

"How's my kid doing?"

"Are the schools succeeding or failing?"

"What works best to help students learn?"

"Do test scores prove the effectiveness of educational programs?"

These are serious questions, and effective accountability systems must address them with clarity and candor. An accountability system that fails to address these common-sense questions does not deserve the support and confidence of citizens or policy-makers.

Simple questions, however, do not necessarily lead to simple (or, more appropriately, simplistic) answers. So it is with educational accountability. These simple questions require complex responses, which I will explore in this article.

## **How's my kid doing?**

A question about the performance of a specific child implies that an effective accountability system will be based on information about individual students rather than groups or classes of students. In other words, an accountability system that can supply information about the progress of a specific child must be based on a series of individual student achievement records. If these records are flawed, then the entire house of accountability is built upon a shaky foundation.

One of the most important developments in educational accountability has been the "value-added" methodology developed by Professor William Sanders and his colleagues at the University of Tennessee (1998). The core of the system is a "student-to-student" comparison. While such a comparison makes common sense, it is rare. By far the majority of state and local accountability systems compare one year's class of students to the previous year's class of students - a comparison that involves almost entirely different individual students.

Such group comparisons never address the fundamental question, "How is my kid doing?" Rather, parents and teachers are given the curious information that their eighth-grade students are better or worse than last year's eighth-grade students. Group comparisons tell us nothing about the progress and educational needs of individual students.

This has serious implications for accountability systems that claim to evaluate educational quality, but ignore information about classroom teaching and learning. An accountability system that contains test scores alone, without the context of additional accountability information about teaching practices and curricula, is incomplete.

A school district that bases its accountability system on test scores alone is analogous to a physician who evaluates physical health based upon indicators such as body temperature or blood pressure, but ignores the other medical indicators that any reasonable physician would regard as essential to a competent diagnosis. In the most bizarre cases, accountability systems not only fail to evaluate substantial parts of the school curriculum, but actually encourage behaviors contrary to those endorsed by the designers of the accountability system.

If parents or stakeholders want an honest and accurate response to the question of "how's my kid doing?" then school systems and districts must use tests or other assessments that will yield the information needed to answer this question. In other words, children need to be assessed to make sure they meet the standards, and only this information will determine how they're "doing."

The most virulent critic of public education would not attend an athletic event and, desiring to know the outcome, settle for a description of the weather and field conditions. "What the heck happened during the GAME?" the critic would demand. In the context of the classroom, parents and students must learn about the process and results of teaching and learning.

If the education "game" is to be taken at least as seriously as many people take their weekend athletic events, then it is reasonable to expect that the data used to evaluate the result should be related to the issue at hand: student and school performance. Thus, if we want to know how well students write, then we must ask them to write. If we want to know whether they can use the scientific method, then we should ask them to design an experiment and draw inferences from a set of data. If we wish to know whether students understand mathematics, we should ask them to use mathematics to solve real problems.

### **Are schools succeeding or failing?**

The second question raised in the name of common sense is built upon the first. Schools succeed only where students succeed. Thus, assessments of schools such as accreditation or typical accountability reports, are only as effective as their evaluation of students. Accountability systems that look only at process and effort will reward a fixation on meetings, plans and strategies while ignoring results. Every initiative, including those that I have advocated - high standards, effective assessments and rigorous accountability - is only effective when it is built upon a foundation that soundly evaluates student achievement.

Mike Schmoker (1999, 1999a) is a leading advocate of the focus on results. He has endured heaps of abuse from those who find process a perfectly suitable substitute for student achievement. While a focus on results is important, the "results" that effective accountability systems must consider extend far beyond test scores alone.

This is not merely a debate over "process vs. results" In fact, a comprehensive accountability system must focus both on measurable elements of the process (specific instructional, assessment and engagement strategies), and on results (indicators of student achievement). Only with such a comprehensive view can we gain some insight into what the adults in the system can do to influence results achieved by the students.

Although the evaluation of students is the foundation of a sound accountability system, an effective accountability system must base its examination of student achievement on more than test scores. Indeed, the fallacy of reporting school-wide success or failure based on single test scores has been widely documented (Bracey 1999). Teaching and learning are multiple interactive

processes, the results of which are much too complex to be captured by a single score.

Accountability systems that depend solely on test scores offer predetermined results: students who are capable test takers will make a school look like a success; students who are not capable test takers will give their school the label of failure. In neither case do the organization, leadership, teaching and educational practices of the school receive a meaningful evaluation.

Instead, effective evaluation of the success of a school can only be measured with multiple information sources over an extended period of time. There must be several indicators that measure the performance not only of students, but also of the adult decision-makers. Such an analysis includes a consideration of resources, teaching methods and student support. Without a consideration of all of these factors, we are left with the ludicrous notion of evaluating the performance of an entire institution based on the test scores of one group of eight-year-olds during one stress-filled, springtime afternoon.

Effective, comprehensive accountability systems distinguish between student achievement and school achievement without losing sight of the fact that the latter is integrally related to the former. For example, state and district test scores might provide some indicators of student achievement. But these results are only meaningful indicators of school achievement when they are placed in the context of the specific educational strategies used by the schools. The structure of such a comprehensive system acknowledges both the distinction and the importance of both questions.

### **What works best to help students learn?**

The third common-sense question addresses the heart of an effective accountability system. The question of program effectiveness is far more complex than the recitation of test scores representative of most accountability systems. Such a "box score" approach is the educational equivalent of giving the ranking of teams at the end of the season without shedding any light on the strategies that led the teams to succeed or fail.

For the disinterested observer who wishes only a cursory overview, the final rankings may suffice. But those interested in the game would want much more. People whose futures depended upon the success of the team would demand a continuous analysis, not only of scores, but also of the individual elements of strategies and programs that lead to success.

In the context of school accountability, stakeholders need to know which programs succeeded and which ones failed. In a field littered with "reforms" and "new ideas," some rational method of evaluation other than popularity, enthusiasm or cost must be used. An accountability system that shows policy-makers how intervention strategies correlate with student results can go a long way toward providing such essential program evaluation information.

Some of the best practitioners of such systematic evaluation are Robert Slavin and his colleagues at John Hopkins University. In their recent book, "Show Me the Evidence!" (1998), they provide examples of how accountability information can be used to monitor program effectiveness. Perhaps the most astonishing aspect of the book is how few programs have been subject to long-term, systematic accountability and analysis.

In sports, media commentators are expected to report not only scores, but they must also demonstrate a thorough understanding of the nuances of the game. They routinely debate the meaning of various statistics. The television commentator or sports journalist who simply writes, "Those with winning records are better teams than those with losing records" would soon be out of a job.

When it comes to sports, we expect the best commentators to look beyond the data and provide insights based on observation, description and qualitative understandings that extend beyond

numerical explanations. Though the athletic analogy may seem clumsy, I will celebrate the day when analysis of educational accountability data is taken as seriously by the media and the public as analysis of last weekend's sports games. When that day comes, we all may finally know what works best to help students learn.

A comprehensive evaluation of what works best to help students learn must go far beyond mere test scores and include analyses of both how the data reported should be interpreted as well as information about the context in which these results were derived.

### **Balancing qualitative and quantitative information**

In educational accountability, numbers are an important part of the story, but they tell only part of the story. The qualitative dimension of accountability - descriptions, narratives and observations about culture and climate - creates a lens through which the quantitative data must be viewed.

What is the qualitative context of the quantitative data? What are the successes, failures, tragedies, and triumphs of this school that help to explain the story behind the numbers? While no principal in America needs one more report to write, a one-page synopsis of a school's qualitative dimension would add greater context to its test scores. Without such a qualitative context, we are left with the sterility of data that, even when presented with abundant statistical complexity, can leave us wondering, "What really happened in that school?" Without the qualitative dimension, our understanding of the "score" is limited, incomplete and possibly inaccurate.

What sort of qualitative information should be included? Information about the school climate and environment, the triumphs and tragedies of the school year, and descriptions of any significant changes in programs, personnel or performance can all be expressed in narrative form. Such narrative information is not a substitute for quantitative data, but rather, gives citizens and policy-makers context in which to interpret numerical results.

### **Meeting the test of common sense**

An effective accountability system must answer at least four common-sense questions: one about individual student achievement, a second about school performance, a third about ways to help students learn, and a fourth about determining educational effectiveness.

In order to provide useful information about student achievement, an accountability system must be based on clear standards that have been communicated to students, parents, teachers, and other district stakeholders. Both quantitative and qualitative indicators that measure whether or not these standards have been met must become integral parts of the accountability system.

School performance must be based on much more than test scores. Though it is likely to include test data, school performance information must also include how those numbers should be interpreted and the context from which test scores arose. This approach to comprehensive accountability is, to be sure, more challenging than simplistic headlines. It is, however, the only approach that meets the simple test of common sense.