

# **"In God We Trust: All Others Bring Data"**

by

Dale Mann, Ph.D. and Charol Shakeshaft, Ph.D.  
Managing Directors  
Interactive, Inc.

1.0	Introduction .....	2
2.0	Publishers .....	2
3.0	Public Schools.....	5
4.0	Consequences .....	7
5.0	Summary.....	9
	Appendix:Defining Scientific Research .....	10

# "In God We Trust: All Others Bring Data"

by

Dale Mann, Ph.D. and Charol Shakeshaft, Ph.D.  
Managing Directors  
Interactive, Inc.

## 1.0 Introduction

The "No Child Left Behind" law passed in January 2002 and its \$50.3 billion budget request sharply accelerates the required use of student data to visit accountability results on schools and publishers. Schools face the loss of Federal funding coupled with Federal subsidies for parents to take their children elsewhere. Publishers face an accountability future not unlike what the pharmaceutical industry experiences with Federal approval of private drugs. Consider the following headline.

"Sepracor's Stock Plunges on F.D.A. Ruling on Drug:  
Allergy Medicine Will Need More Testing"

[Andrew Pollack, *The New York Times*,  
Friday, March 8, 2002, p D2]

The phrase, "In God we trust: all others bring data", is literally a shibboleth<sup>1</sup> among top officials in the current Department of Education and captures one long-term effect this administration is likely to have on school practice and participants. One measure of their intentions is the 111 times that phrases like "scientifically based research" and "evidence based decisions" occur in the new act. To the same end, the House Committee on Education and the Workforce said that one purpose of H.R. 3801, the Education Sciences Reform Act, was to "put an end to education fads that masquerade as sound science". [Elizabeth B. Guerard, "House bill might kill ed tech research agencies", eSchool News, March 29, 2002.]

This reviews the consequences for publishers and for schools of those parts of the act that deal with student performance data.

---

<sup>1</sup> "Shibboleth" means "password" and was used by Biblical Hebrews to distinguish friends from foes.

## **2.0 Publishers**

Publishers will be the first to point out that market decisions, including whole state textbook adoption procedures, make them accountable in ways that are almost unique among the participants in schooling. If schools don't believe in a product, they don't buy it. But Federal officials intend to add new hurdles to the process and point to the medical model, specifically, drug testing which is paid for by pharmaceutical manufacturers but approved (or not) by the Food and Drug Administration.

“The goal is to follow the lead of the medical profession, in relying more on well-crafted research to guide practice, said Joseph C. Conaty, the acting deputy assistant secretary for the Education Department's office of elementary and secondary education. ‘This is the first, early steps of education moving in this direction’ he said at a meeting this month at the National Research Council.” [“Law Mandates Scientific Base for Research” Lynn Olson and Debra Viadero, *Education Week*, January 30, 2002, v XXI, n 20]

The path that the Federal government wants to block is described by Mary Fulton of the Education Commission of the States: “If publishers can waltz into a district and sell their program based on the sales pitch and the PR, then why do rigorous evaluations?” [*ibid.*, p 15]

The authors quote Grover Whitehurst, the assistant secretary for research, “...(C)ompanies that market educational products also will have to take some responsibility for investing in clinical trials and evaluations. ‘The time will come when it will become a distinct marketing advantage.’” [*ibid.*, p 15]

Publishers in the early childhood market segment will recognize a similar impulse in the recently launched, “Preschool Curriculum Evaluation Research Grant Program” (PCER)---“The purpose of this program is to implement rigorous evaluations of preschool curriculum that will provide information to support informed choices of classroom curricula for early childhood programs.” [US OERI, CFDA No. 84.305J, *Federal Register* v 66, n 242, December 17, 2001] Publishers whose products are documented to be effective will have earned Dr. Whitehurst's “distinct marketing advantage”. PCER is a harbinger although in this case one that is fully funded by the government.

And, the Department intends to set up a “What Works Clearinghouse” to evaluate the claims made about textbooks, software and videos.

Specifically with respect to technology, the Department will evaluate the effectiveness of e-learning as a whole. The law requires a 3-year, \$15 million study of, “The conditions and practices under which educational technology is

effective in increasing student academic achievement, as well as the ability of teachers to integrate technology effectively into curricula and instruction.” The Department’s current technology-related expenditures total \$850 million across 10 programs: they are the likely sites for the first application of any results from the Department’s determination of effectiveness.

The school health risk prevention field offers a precedent for the attempt to impose “evidence-based” criteria on curriculum. Schools are nominally required to demonstrate that the curriculum they purchase with "Safe and Drug-Free Schools" money meets Federally established "Principles of Effectiveness". The “No Child Left Behind” law extends that procedure to any activity financed under the \$1.25 billion<sup>2</sup> "21st Century Community Learning Centers" program. The authorizing language for "Reading First" (\$900 million) and "Early Reading First" (\$75 million) programs restricts expenditures to programs that are "scientific, research-based".

The Public Health Service was concerned that school districts would under-serve children by adopting health risk prevention curriculums that were popular but not proven to be effective. In 1998, the Service promulgated the "Principles of Effectiveness". Districts were required to follow four steps: (1) begin with a needs assessment with objective data; (2) establish measurable goals and objectives with community input; (3) *select effective, research-based programs* [emphases supplied, Authors]; and (4) periodically evaluate those programs.

The next year, a survey of districts reported the following patterns.

**Table 1. Drug Abuse Prevention Programs Used by Public School Districts (%)**

<b>Program</b>	<b>% Using</b>
D.A.R.E	83
Here’s looking at you	63
McGruff’s drug prevention	52
Life Skills Training	41
Project ALERT	31
Learning to live drug free	27
Discover: skills for life	26

Source: Denise Halfors, "Policy Changes in the Safe and Drug Free Schools Program", University of North Carolina, School of Public Health (September 99)

To that point, D.A.R.E. had been unable empirically to establish its positive effects with children but it was still the runaway best seller. The McGruff program, number 3 on the popularity list, was no longer being published and the

<sup>2</sup> Dollars are authorization not appropriation.

two programs with the most “scientifically based research” documenting their positive outcomes for children were used less than half as frequently D.A.R.E.

In September 2000, not a single studied district had a list of research-based programs from among which to select. The Public Health Service delegated enforcing district adherence to *The Principles of Effectiveness* to the states and Hallfor's research indicates that none of the 10 study states had any intention to sanction districts that did not follow *The Principles*.

The ‘programs that work’ label was originally from the Centers for Disease Control of the Public Health Service and until recently only two programs were on the recommended list. In 2001, two more agencies began to identify and promote officially certified programs--one from the US Department of Health and Human Services and one from the US Department of Education. The PHS Center for Substance Abuse Prevention set up a review process for health education programs. Sixty-two programs were expected to be identified as “model programs” by the National Registry of Prevention Programs by the end of 2001.

In January 2001, the US Department of Education announced that it had selected nine programs as Safe and Drug-Free Schools “exemplary” programs with another 33 additional programs described as “promising”. (A total of 132 programs were reviewed.) Criteria for exemplary status included “quality, effectiveness, usefulness to others and educational significance”.

Ted Sanders, a former Federal official and chief state school officer who is now the president of the Education Commission of the States observed that state departments of education “...are often under intense political pressure to recognize certain programs or products as meeting standards of evidence, whether they do or not.” [Olson and Viadero, *Ibid.*] And, *Education Week* continues, “...superintendents, principals, and teachers...may be too immersed in their day-to-day challenges to, in effect, research the research.” [*ibid.*]

No one should be surprised or misled by the political reality of school decision-making or by the contentious process of research. Local pushback is as inevitable as is a stretched out schedule for implementation (witness, the 12 years allowed in “No Child Left Behind” to get all students up to standards). The process is in fact an opportunity that prudent businesses will use to equip themselves for new demands.

The current initiatives have broad support among legislators and administrators; holding publishers to high standards is political high ground; and these initiatives are the next turn of the policy wheel. To paraphrase the country western song, publishers have discovered, “If you want to sell in Texas, you’ve got to have TAAS in the brochure”. Just as consumers look for EPA ratings on refrigerators,

risk averse school administrators will reduce their vulnerability in making curriculum adoption decisions by citing stamps of Federal education approval for otherwise local decisions.

### **3.0 Public Schools**

Three of the major sections of No Child Left Behind deal directly with student performance data: (1) annual testing; (2) academic improvement (measured by test gains); and (3) report cards. "Secretary Paige has indicated that the department favors the development of student data-management systems and online testing, which would help give educators immediate access to data about their students' academic strengths and weaknesses." 'Anne Wujcik, "HR Analysis", *The Heller Report Educational Technology Markets*, March 2002, v 13, n5. Taken together, the procedures may create a giant "naturally occurring experiment" in which we will be able to associate changes in student outcomes with changes in various inputs, including curriculum.

#### **3.1 Annual Testing: Every Grade, 3-8, Reading And Math**

States have three years, beginning next September to either design or adapt tests that will be used, statewide, to measure the performance of all students in grades three through eight in reading and mathematics<sup>3</sup>. The Federal government will pay part of the cost of developing and/or implementing these tests. The tests can be state standards-based<sup>4</sup> or normed (e.g., the Stanford 9 or the Iowa Tests of Basic Skills). The former are more likely to correspond to the delivered curriculum in that state's classrooms while the latter are more convenient.

Second, states are required to make a sample of their 4th and 8th grade students available for testing by the National Assessment of Educational Progress. Comparing the national, NAEP results to the local test results is intended *inter alia* to embarrass jurisdictions that are tempted to guarantee their (nominal) success by administering easy tests. To shut off this "What year was the War of 1812?" gambit, the Federal government will pay for 100% of NAEP data collection costs. Additionally, the Federal government will monitor progress and require that goals are increased at least every three years.

---

<sup>3</sup> The 1994 act required statewide reading and math tests for at least one grade each in elementary, middle and secondary schools. The intensity of the newly mandated test schedule is one measure of the increased pressure for data-driven decisions (and consequences). The other measure is the coupling of sanctions for adults to performance of children. Upper grades testing is sure to follow.

<sup>4</sup> Nine states already have tests for reading and math in grades 3-8 that are designed to measure progress on that state's standards. (*Education Week*, January 9, 2002)

In 2007-08, states must add science tests for at least one grade each in elementary, middle and secondary schools.

"No Child Left Behind" requires that each state publish 'report cards' detailing student achievement data by district and by subgroup. All test results must be reported by groups including race, income and other categories. Without this "disaggregation" feature, schools and districts would continue to mask their lack of progress with particular groups by the simple device of reporting whole group, summarized or aggregated data. District-level reports must be similar and therefore comparable.

### **3.2 Academic Improvement**

This is the source of the new, "adequate yearly progress" yardstick. States are required to attain "academic proficiency" for all students but: (a) they get to define "academic proficiency"; and (b) they have 12 years to achieve the "all students" goal. (c) Schools where students increase their proficiency but fall short of "adequate yearly progress", will be in a "safe harbor"---temporarily exempted from the sanctions otherwise triggered by failure.

The consequences for failure kick in over five years. At the end of two years of inadequate performance by a school, districts are required to provide "technical assistance" and, more ominously, a public school choice, "opt-out-of-failure" opportunity for parents including free transportation to the new school.

At the end of three years of inadequate performance, the school must provide supplemental services---including private tutoring---that the parents choose.

After four consecutive years of inadequate performance, districts are required to replace staff members, adopt new curriculums, etc.

If five years pass without sufficient improvement<sup>5</sup>, the law anticipates that schools will be closed and taken over by the state or re-opened as charter schools.

Some of the features just described dilute the impact of the act. That predictable feature of legislative coalition building made possible a Senate vote of 87-10 and a House vote of 281-41. Such temporary compromises should not be misinterpreted. The provision of health care in the United States has been profoundly and irrevocably transformed by exactly the kind of public, incremental data-driven processes that are now underway in public schooling. There will only be more scrutiny of data, more consequences for different aspects of student and

---

<sup>5</sup> Scores of New York City elementary schools, officially designated as failing, have been on those lists for decades. Compared to current practice, the five-year warning-to-takeover cycle is precipitate.

adult performance and therefore more opportunities for services related to data. Many school administrators will have a new interest in data, all are certain to follow.

### **3.3 The Prospect for Implementation**

Essentially, the act makes states responsible for districts, and districts responsible for schools. But until the Department issues regulations, we will not know exactly what states are to do and what districts are to do. The Department of Education is currently managing the preliminary review of its regulations with a final release scheduled for August. (Government watchers will be bemused to note that the states must submit their comprehensive testing plans before the Federal government tells them what is required of them.) Those regulations will be followed by State and local implementation of the guidelines and by program funding.

The unevenness of the process can be measured by the fact that the 1994 re-authorization of ESEA included core requirements for standards and testing which have still not been fully implemented in the states and 14,000 school districts (see *Education Week*, November 28, 2001). High academic standards, assessments aligned to the standards and accountability systems to force compliance are not uniformly in place. Getting school practice to conform to "No Child Left Behind" will be a similar years long process and thus a public and private leadership opportunity spread over months and years.

## **4.0 Consequences**

### **4.1 Consequences for School Administrators**

Now that the federal government will force local districts to collect data and especially to report disaggregated data, school people will be more interested than ever in student performance data, although the majority have neither the background nor the time to do the required analyses in-house.

School administrators and school boards will be intensely interested in data, how they are collected, what they are alleged to measure, how they are analyzed and interpreted. Until "No Child Left Behind", those were psychometric discussions for academics. Now they are financial discussions for people who want to maintain their jobs, their control, their institutions. Consider the following table.

**Table 2: Take-Over Districts by Year and Enrollment**

District	Takeover Year	Enrollment
Jersey City, NJ	1989	31,000
Boston, MA	1991	63,000
Paterson, NJ	1991	25,000
Compton, CA	1993	33,000
Chester-Upland, PA	1994	6,000
Chicago, IL	1995	431,000
Newark, NJ	1995	41,000
Cleveland, OH	1995	75,000
District of Columbia	1995	68,000
Baltimore, MD	1997	96,000
Hartford, CT	1997	68,000
Detroit, MI	1999	168,000
Philadelphia, PA	2002	200,000

Nine states and the District of Columbia have already superseded a dozen major districts enrolling more than a million children. In the future, the cross-group, federally-required comparative spotlight will make inaction in the absence of “adequate yearly progress” harder to justify. Of America’s 80,000 school buildings, 9,000 are believed to be on a “watch list” for federal/state action.

**4.2 Consequences for Parents and Communities**

Think of the high visibility controversies for police that have followed the disclosure of racial profiling. The requirement that school data reports be broken down by subgroups, ("all girls", "all Latino 14 year olds at Jefferson Middle School", "all 13 year olds two or more grade levels behind in reading") will likely arm public and parent involvement as never before. As long as schools or districts could report only whole school, whole district performance, parents of low achieving children had no alternative but private guilt about their child's departure from a “norm”. They were not likely to understand how many other parents shared their anguish. Disaggregating data by income, language of origin, gender, etc. is certain to provide a crystal-clear basis for political organization and therefore pressure to improve that has been absent in schooling.

**4.3 Consequences for Teachers**

The availability of finely grained student performance data will prompt attention to questions of teacher effectiveness as well as to curriculum effectiveness. With respect to teachers, appropriately analyzed data make it possible to track the cumulative effects on students of a succession of teachers; for example, the retarding effect on student learning of a succession of low-performing teachers. Whether the prospect of clarifying such matters is a bane or a boon will vary by

distance from the classroom. The National Education Association observes that, "Consumers demand more information from the labels on a can of soup than the current law does in passing judgment on students and schools" but the final commitment to truth in teaching is not yet clear. Defended by their organizations, teachers are likely to be able to resist the career-determining versions of accountability for as long as another decade.

## **5.0 Summary**

Hoping that better data will promote better decisions is an old family tradition among Washington policy makers. The Federal R&D centers from the Kennedy Administration and the national labs from the Johnson Administration were infrastructure responses to the impulse for rational decision-making which continues to be evident. The old National Institute of Education and the National Diffusion Network were arrows in the rational quiver. Secretary of Education Ted Bell raised the stakes with his infamous "Wall Chart" ranking of the American states by student achievement. His successor, Bill Bennett provoked additional attention to related issues with his "What Works" series.

In the foreseeable future, holding publishers and schools to "high standards for children" is political high ground. Like the "standards" movement, it has the twin virtues of being easy to understand and cheap (if vexed) to implement.

The provisions of "No Child Left Behind" are a guarantee of full-time employment for the members of the American Education Research Association who, along with threatened administrators and teachers will debate endlessly the meaning of "scientific evidence" (for the language of the act, see below). David Kearns, the president and chief executive officer of Xerox and an undersecretary of education famously observed that "Education is the only endeavor I know in which, if you do good, nothing happens to you. And if you do bad, nothing happens to you." Educators could chuckle indifferently at Kearns because no one had good data and, even more to the point, there were not clear connections between student performance data and various dimensions of the school experience. The data requirements of "No Child Left Behind" create new sanctions and especially new, wholesome possibilities.

## **Appendix: Defining Scientific Research**

*Following is the definition of “scientifically based research” as it appears in the “No Child Left Behind” Act of 2001.*

“The term ‘scientifically based research’ (A) means research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs; and (B) includes research that:

- employs systematic, empirical methods that draw on observation or experiment;
- involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;
- relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators;
- is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls;
- ensures that experimental studies are presented in sufficient detail and clarity to allow for replication or, at a minimum, offer the opportunity to build systematically on their findings; and
- has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review.”

*Education Week, Volume XXI, Number 20. January 30, 2002. p.14.*

### **About the Authors**

Dale Mann, Ph.D. and Charol Shakeshaft, Ph.D. are Managing Directors of Interactive, Inc. Dale Mann ([dmann@interactiveinc.org](mailto:dmann@interactiveinc.org)) is also professor emeritus, Teachers College, Columbia University. Charol Shakeshaft ([CharolSakeshaft@interactiveinc.org](mailto:CharolSakeshaft@interactiveinc.org)) is also professor, Foundations, Leadership and Policy Studies, Hofstra University.

Interactive, Inc.  
11 Stewart Avenue  
Huntington, NY 11743  
(631) 351-1190

Interactive, Inc.