Introduction and Overview

Williamson M. Evers and Herbert J. Walberg

Although educators and school boards sometimes resist the idea, accountability is sorely needed in America's schools. This country's children are as able as children in other affluent, democratic nations, but they fall farther behind the longer they are in school, notwithstanding nearly the highest per-student costs.[1] In most instances, they make the least progress and wind up at or near the bottom of international rankings of achievement as they finish high school. Furthermore, a quarter century of numerous reforms, substantially more money, and rising student abilities has generally failed to raise achievement test scores.

Americans take great pride in the superior and ever-increasing effectiveness and efficiency of most of our industries. Yet our schools fall behind those in other countries and have become less rather than more efficient, which is far from what we would want, given their central importance in the American economy and society. School productivity or the relation of achievement to costs was 65 percent

[1]Herbert J. Walberg, "Achievement in American Schools," in *A Primer on America's Schools*, Terry M. Moe, editor, Stanford, CA: Hoover Institution Press, 2001.

higher in 1970–71 than in 1998–99.[2] Surveys of students, citizens, and employers, moreover, reveal substantial dissatisfaction with American schools.

If schools were doing well or even passably well, policymakers might be deservedly reluctant to insist on substantially greater accountability. Since schools are not doing well, however, the burden of proof in explaining this state of affairs should be on the present system. Yet the American K–12 system is distinctively unaccountable compared with other aspects of American life and compared with education systems in other countries. In the work world, for example, management and employees are held accountable. Those that do well gain merit raises, but in other cases, heads roll. The performance of sports and entertainment figures is closely measured, ranked, and encouraged through a variety of incentives. Firms, workers, athletes, and entertainers compete, and the marketplace creates incentives for their efficient performance, holds them closely accountable, and rewards success. The accountability principle extends to most sectors of American life; wise parents hold their children accountable for their behavior, and part of growing up is learning how to be self-accountable for one's ethics, ideals, and goals.

[2]Caroline M. Hoxby, "School Choice and School Productivity: (Or, Could School Choice be a Tide that Lifts All Boats?)," Washington, DC: National Bureau of Economic Research Conference on the Economics of School Choice, 2001. Professor Hoxby concluded as follows: "Consider the simplest productivity calculation, achievement per dollar, without any attempt to control for student characteristics. Such a calculation (which I describe in detail below) suggests that average public school productivity was about 65 percent higher in 1970–71 than in 1998–99. [If we] were simply to restore school productivity to its 1970–71 level, then the average student in the United States would be scoring at an advanced level where fewer than ten percent of students now score" (p. 2).

Contrary to many claims, higher education also shows poor achievement productivity. Alexander W. Astin, in "Undergraduate Achievement and Institutional Excellence" (*Science,* 1968, 161, 661–668), much extended in his *What Matters in College* (San Francisco, CA: Jossey-Bass, 1993), showed that wide variations in spending, facilities, library holdings, and similar costly inputs have no effect on value-added college and university achievement.

Compared to schools in other countries, those in the United States are subject to little accountability. The U.S. school system lacks the marketplace accountability of schools competing with one another and the further accountability of large-scale examination systems, both of which are associated with higher achievement.[3]

Educators' resistance to testing and accountability is not surprising and is rather to be expected. The essayist and playwright George Bernard Shaw said all professions are conspiracies against lay people. Neither doctors nor bricklayers would choose to be accountable for their efforts—it would be much easier for them to say that all's well as they request higher compensation.

Professional associations and unions have their members' interests and welfare as their first priority. So dealing with them in free society requires clients or consumers to represent and press for their own interests. In the case of schools, it is not unions but legislators, school boards, education leaders, taxpayers, and parents that have been remiss in failing to acquire information on the standings and efficiency of states, school districts, schools, and individual staff members.

Oddly, many educational psychologists and specialists in educational testing, evaluation, and statistics have been silent or reticent on the need for testing and accountability. Many, however, are in schools of education where views are unlike those in other walks of life. A recent Public Agenda survey of education professors showed that 64 percent think schools should avoid competition.[4] Education professors also differ from employers and other professionals on

[3]Ludger Woessmann, "Why Students in Some Countries Do Better," *Education Matters,* Summer 2001, 65–74.

[4]See Walberg, 2001. Nearly two-thirds of teacher educators admitted that education programs often fail to prepare candidates for teaching in the real world, and only 4 percent reported that their programs typically dismiss students found unsuitable for teaching. Thus, even starting with their undergraduate education, many prospective educators are laden with anticompetitive ideas against standards and incentives.

measuring standards or even employing them at all. Employers employ standardized examinations for hiring. So do selective colleges and graduate and professional schools for admission decisions. Such examinations are required in law, medicine, and other fields for licensing, in part because they are objective and reliable. Yet 78 percent of teacher educators wanted less reliance on objective examinations.

Even National Academy of Sciences (NAS) education committees, heavily populated by educators and education professors, have spoken out against tests and procedures designed to elevate educational standards and raise the stakes for success and failure. A recent NAS report, for example, warned against using scores from a single assessment for promoting students in school and retaining them in a grade for an additional year.[5]

The NAS committee failed, however, to point out adequately the advantages of relying on a single assessment. Using a single test is clear and definitive. It presents a challenge, which students and teachers have a strong incentive to meet. It has successful precedents in many fields, including K–12 schools outside the United States. Meeting standards in school is preparation for college, employment, sports, and other common endeavors where clear, objective, unambiguous standards are routine.

The NAS committee also neglected to point out adequately the harm that school systems, particularly those in big cities, have done to the careers of poor youth, who have not learned basic knowledge and skills, and the unfairness imposed on students capable of meeting higher standards, who have graduated with devalued diplomas because schools have not demanded high levels of achievement and have not tested to make sure those high levels have been attained.

---

[5]Jay P. Heubert and Robert M. Hauser, editors, *High Stakes: Testing for Tracking, Promotion, and Graduation,* Washington, DC: National Academy of Sciences, 1999. Fifteen of 25 members of the committee were school people and education professors.

The NAS committee held that tests should meet various stringent requirements before being used as criteria for high school graduation. Following their recommendations would be expensive, difficult, time-consuming, and unprecedented. Few states, perhaps none, have perfectly aligned their goals, standards, teaching materials, and tests with one another. Even if they did, frequent changes in one aspect may require changes in other aspects, making it extremely difficult to mount and maintain a system acceptable to the NAS committee.

Yet examinations with only "low stakes" consequences would mean a continuing escape from accountability for failing students, schools, and systems. There is justification, of course, for phasing in and elevating standards in measured steps, but there is little reason for retaining the present unproductive status quo until a perfect system can be devised.[6]

After a quarter-century of poor progress in educational productivity, the burden of proof is on schools rather than on tests per se or on the idea of accountability, and the time for inaugurating high standards and accountability is now. As common sense would suggest, moreover, research on standards and accountability shows their beneficial effects even in the many places where they have been put into effect in a far from ideal way.

Accountability critics, for example, maintained that strict Standards of Learning imposed on Virginia schools for accreditation would be a debacle. Instead, the percentage of schools meeting the standards rose from 2 percent to 40 percent from 1999 to 2001, and many more schools are expected soon to achieve the standards. Said State Superintendent of Education

[6]Without qualification, perfect alignment of lessons and test might mean "teaching to the test," which may risk trivializing both teaching and testing. Good tests usually rely on sampling of content, just as surveys rely on a small sample to estimate population views. They usually also avoid test items exactly like those in a textbook or class lesson so that students must master the application of principles to new or less familiar examples.

Jo Lynne DeMary, "In more and more schools, teachers and administrators are analyzing curricula and making the changes needed to improve instruction and increase student achievement."[7] Scholars have documented other impressive success stories.[8]

OVERVIEW

To analyze accountability issues, we have gathered a balanced set of contributions from authors who are specialists in studying education from the perspectives of the social and behavioral sciences: history, economics, political science, and psychology. These perspectives, sometimes combined in a single article, should enable readers to gain an understanding of what is known about accountability, what still needs to be learned, what should be done, and what is best avoided in devising accountability systems.

In the opening chapter, historian Diane Ravitch distinguishes policymakers' interest in results as measured by tests and reinforced by accountability mechanisms with professional educators' interest in improvement—possibly using tests for diagnostic purposes but not for accountability. The professional expects that spending more money in ways that professional educators like will yield improvement. Ravitch traces this view to the turn of the twentieth century when testing began to be used in college admissions. She carries her story through the Progressive era with its scientific and managerial ethos, in which testing was used in the hope of

[7]Alan Richard, "More Virginia Schools Hit Mark on Exams Used for Ratings," *Education Week,* October 24, 2001, pp. 26, 30.

[8]See subsequent chapters in this volume and Diane Ravitch (Editor), *Brookings Papers on Education Policy,* Washington, DC: Brookings Institution, 2001, particularly the chapters by Julian Betts and Robert Costrell; Herbert Walberg; Chester Finn and Marci Kanstoroom; David Grissmer and Ann Flanagan; and John Bishop, Ferran Mane, Michael Bishop, and Joan Moriarity, which summarize evidence that states that set clear standards, align curricula, publicize comparative results, and have begun incentive systems have made better than average achievement progress.

improving teaching practices and with the intent of fending off scrutiny from parents and taxpayers.

Policy analyst Chester E. Finn, Jr., uses the metaphor of the variegated pairings in the 1969 motion picture *Bob & Carol & Ted & Alice* to illuminate how choice and accountability work together. He finds that the best mixture combines choice (which he nicknames "Alice") among rival institutions, such as charter schools and voucher schools, with academic standards and external testing ("Ted"). Along the way, he reveals the flaws of other possible matches and mismatches of approaches.

For many years, economist Caroline Hoxby was a test skeptic because she thought that the benefits from testing (in the absence of accountability incentives) were too small to justify the cost. She has changed her mind. She finds that she had underestimated how much schools would improve to avoid being exposed as low performers, and she now realizes how inexpensive testing is. Hoxby calculates what testing actually costs, and she finds it to be one of the most cost-effective of all school reforms.

Economist Eric A. Hanushek and political scientist Margaret E. Raymond focus on incentives. They point out that, at present, most consequences fall on students; schools and teachers feel consequences only indirectly. Hanushek and Raymond describe the components of an effective accountability system. It should pinpoint where problems are and encourage appropriate change. They find, however, that knowledge of what works best in the classroom is still inadequate.

Policy analyst Lance T. Izumi and political scientist Williamson M. Evers analyze accountability in three states with comparatively strong systems: California, Texas, and Florida. Concentrating on a few states allows the authors to depict how accountability works in practice. They are most hopeful about the Florida system with its standards-based test, school report cards, merit pay for teachers, and exit vouchers for students in failing schools.

Psychologist Herbert J. Walberg points to the features needed in effective accountability systems and provides examples of consumer-friendly reporting of accountability results from actual accountability systems. Walberg contends that some of the existing tests and accountability systems are good enough to do the job and warns us not to let a yearning for perfection block implementation of approaches that will in fact work better than existing systems and continue to improve with experience.

These contributions from education policy specialists help us see how we came to have failing schools, low-performing students, and little accountability. They show how we can devise the affordable, reasonable, and workable accountability systems and incentives we need to raise student learning. Myths about accountability have misled some fair-minded people into fearing that accountability will somehow discourage learning. But in actuality, accountability and appropriate incentives offer our best hope for improving American public schools.