

INTRODUCTION

Critical Concerns About Accountability Concepts and Practices

Kenneth A. Sirotnik

January 9, 2003

Dear Ken,

I was visiting Susan Ohanian's web site after our talk, and I saw the following quote: "It is the greatest of all mistakes to do nothing because you can only do a little. Do what you can." (Sydney Smith)

I have reconsidered taking that other job in the private sector after reading that quote. Unless you happen to be in a classroom at this time, it is impossible to describe how awful the environment is. The testing mania is so pervasive that I feel that my contribution in the next few decades will be minimal, if not zero. Should I continue in education, knowing that my ability to contribute will be negligible?

Another thing that concerns me is that in the current environment, the worst teachers are the ones who appear to be doing the best job. Those who refuse to teach to the test are seen as the worst teachers, because their students won't be prepared to score highly. Do you think teachers such as me will remain in education, given the current climate? Do you know how frustrating it is to know that you are doing a better job than everyone else, but that they are getting evaluated as a better teacher?

The teacher I am replacing has got to be the worst teacher I have

ever seen. I had a chance to observe her in class for a week or so, and I was horrified. She knows little math, and is unable to help students appreciate the applications and beauty of math, and so students resort to memorization to pass her moronic tests. The students are so screwed up mathematically that even being here two months, I am still confused as to the proper course to follow. My conscience cannot allow me to imitate her method, even though it is clear that the students and the administrators wish that I would. Most of the students think she was a great teacher. I know why. She expects little, tells them exactly what will be on the test, and they all get great grades, even though they learn absolutely nothing. She is the prototypical bad teacher that we read about, and the one we always say should not be in the classroom.

But here's the irony, Ken. The administration loves her! The parents love her! The students love her! The administration loves her, because she keeps the test scores high. The parents love her because the students are happy. The students love her because it's easy to get a good grade. But the students learn nothing. (There are some exceptions. Some of the students resent the poor teaching, and know that they are screwed mathematically when they go to college. They are glad that I came.)

There are other teachers in this high school that are equally awful and they do a terrible job. They are blissfully unaware of their incompetence, so they can be excused. But many teachers out there know they are doing a lousy job, but simply don't care. They are the real traitors to education. I had some of the students from this school when I taught at the University, and I am beginning to understand why they were so poorly prepared. One of the students in my calculus class passed the placement test, but could not add fractions!

When I ask the students to participate and discover at least some things on their own, or ask them to solve challenging problems, they are confused and angry. They want me to spoon feed them, and ask me for things to memorize so that they can get a good grade on the tests. The superintendent is furious with me, and the principal has asked me to concentrate on preparing the students for the tests, which I refuse to do. The parents are next, and I am the subject of discussion in the community. Some have heard of my methods and want me fired.

When I came to the UW in Seattle, I sacrificed a great deal. I turned my whole life upside down. My family was separated from me for a time, my financial situation was affected, and has still not recovered, and I lost 3 or four years of my career. Now I find that, after having made monumental efforts to improve my teaching skills, that my services are much less valuable now! I am doing a better job, but I am judged to be a poor teacher. Those who are doing an awful job are judged to be doing an excellent job.

How can I go on knowing that doing a good job will cause me to be

Introduction

evaluated poorly? How can I watch other teachers do a terrible job and get excellent evaluations? Where does that leave me? How can anything ever change when the people who evaluate me are unable to recognize a good teacher? The administration here wouldn't know a good teacher if it came up and bit them on the ass. How am I expected to do a good job if I am not recognized for doing so? Is that humanly possible for an extended period of time?

One answer is that I can prepare other teachers. But if I reach them in my classes, and they understand what doing a good job really is, aren't I condemning them to being treated as I currently am? Would it not be better to tell them the truth, that they are likely to be evaluated as excellent if they teach "Parrot Math?" I can not and will not do that. So what is left? And think about George W's Leave No Child Behind act. We know the testing mania is going to get worse. How can I prepare good teachers in this climate?

If you think about the situation long enough, you realize the only answer. The only logical way to proceed is to pretend like you are teaching parrot math, make sure the students are prepared for those idiotic tests, but then surreptitiously try and do some good when no one is watching. In other words, like the opening quote, I can do a little. But knowing that I am being prevented from doing a lot is frustrating me beyond belief. Can my conscience allow me to do a lousy job most of the time, just so that I can do a good job some of the time? Such a superhuman effort is beyond me. I don't think that I can do it. Can anyone?

Either cave in and teach to the test, or rebel and raise hell and cause yourself a lot of trouble. Those who do the former are rewarded. Is it any wonder why they do? Take the easy road, and get rewarded. Take the high road and get scorned and humiliated. Is that what being an educator is all about, Ken? Isn't there any more than that to look forward to? Should I spend the next 30 years of my life in an uphill battle that I cannot win, when I have so many other opportunities? Is it worth the headaches? I don't think that I'm strong enough to fake it just so that I can stay around just to do a little bit of good. I am at the end of my rope. Franklin Roosevelt said that when you get to the end of your rope, tie a knot and hang on. Well, I am hanging by a very small knot, and I'm slipping. I wish I had better news for you.

Convince me to stay in education, Ken. Or tell me it's time to leave. I need some advice, and I don't know a single person whose advice I value more than yours. If you feel like telling me off for unloading this on you, go ahead. I probably deserve it. If you want to tell me to buck up and quit whining, do that. Or tell me that education needs people like me. Just tell me the truth.

Andy

Andy is a real person and this is a real letter. He is one of the graduate students I have worked with, and he has moved back to his mid-western home state. He is an experienced high school and college math teacher. He is a good teacher. What he says about his experience as an educator in the present era of high-stakes testing and accountability is compelling and chilling. It is an expression of both hope and despair. And I don't think his story is at all unique.

In a way, this book is dedicated to Andy and all the good educators like him who are struggling to do their best in an increasingly frustrating policy environment. And although critical, this book is intended to be more about hope than despair. In this chapter, I offer a brief perspective on the problems and pitfalls of accountability, and then share some important claims that I think can be defended and that constitute the foci for subsequent chapters. In the final chapter, I return to critical analysis but also to alternative perspectives that offer hope for more responsible practices.

ACCOUNTABILITY AND ITS PROBLEMS

Teaching is tough under any circumstance, and there will always be better and worse teachers and better and worse conditions and circumstances in which to be teaching. It is unfortunate, indeed, as in Andy's case, when a tough job is made even tougher by a policy context seemingly more bent on threat and punishment than on encouragement and real support for educational improvement. Although certainly debatable, it is my belief that this pretty much characterizes the current and past contexts of high-stakes accountability as we begin the new millennium.

Evidence continues to mount about teacher and principal demoralization and attrition over frustration about the effects of mandated testing for high-stakes accountability (Goodnough, 2001; Jones, Hardin, Chapman, Yarbrough, & Davis, 1999; McNeil, 2000; Whitford & Jones, 2000; Winerip, 2003). Some are getting angry enough to become activists in boycotting tests and even in releasing test items (e.g., "Chicago Teachers," 2002; Gehring, 2002). Some are resorting to cheating (e.g., Hoff, 2000; Keller, 2002; Sandham, 1999). And, tragically, there may be even more horrible outcomes.¹

Negative reactions are not limited just to educators. All across the nation, students and parents have gotten into the fray (Manzo, 2001; Schrag, 2000). The recent Phi Delta Kappan/Gallup poll (Rose & Gallup,

Introduction

2003) clearly indicates the skepticism of parents toward standardized tests and their utility in judging their children and their children's schools for high-stakes accountability. And legal challenges have begun or are underway in several states to counter the fallout from high-stakes testing and accountability practices. Notwithstanding the U.S. Supreme Court's reluctance thus far to consider cases against high-stakes exit examinations, for example, the Louisiana case (Walsh, 2002), momentum against the use of tests as gatekeepers is mounting and may eventually lead to a successful challenge. Allowing the use of basic-skills (or minimum competency) tests as gatekeepers (e.g., for high school graduation) is one thing; but when substantial numbers of middle- and upper-class students start failing high-standards exit exams, the protests and legal challenges are likely to pick up a real head of steam.

No wonder some states are already backing off and even contemplating how to make it easier to pass their mandated tests (Bowman, 2001; Dillon, 2003; Schemo, 2002). But the real injustice, ironically, is to the very students high-stakes accountability advocates claim to be concerned most about—economically poor students and students of color (Orfield & Kornhaber, 2001). These students are clearly the victims of the fallout of high-stakes accountability practices in disproportionately higher numbers (Jacob, 2001).

And what is the fallout? Consider the less-than-encouraging studies of the "Texas miracle" and the long-term efforts in Kentucky. According to the extensive analysis of the available data by Haney (2000), "Texas schools are devoting a huge amount of time and energy to preparing students specifically for TAAS [Texas Assessment of Academic Skills]...Emphasis on TAAS is hurting more than helping teaching and learning ... Emphasis on TAAS is particularly harmful to at-risk students ... Emphasis on TAAS contributes to retention in grade and dropping out of school" (sect. 8, p. 6.). The analysis by McNeil (2000) essentially reports very similar findings and conclusions.

The story in Kentucky is also similar. According to Whitford and Jones (2000), "Kentucky's accountability approach has undermined the very changes in teaching and learning that it was intended to promote, calling into question the use of performance assessment for high-stakes accountability" (p. 21).

Sadly, all this fallout is predictable and has been well-known long before the current incarnation of high-stakes accountability. We've had lots of experience in this country over the past several decades and more of test-driven accountability schemes designed to hold educators' and students' feet to fire. A good summary of the problems and pitfalls is

provided by Heubert and Hauser (1999). Specific issues are addressed by Shepard (1991) on the effects of high-stakes testing on classroom instruction, by Shepard and Smith (1989) on the effects of flunking students in their formative development as learners, and by Nolan, Haladyna, and Hass (1992) on troubling issues more generally with achievement tests. Buly and Valencia (2002) have noted recently the particularly pernicious effects of unwarranted and oversimplistic assumptions about the meaning of setting performance standards on high-stakes, on-demand tests. In their comprehensive analysis of the patterns of failure for those kids "below the bar," they conclude:

No single measure or intervention can possibly meet the needs of all, or even most, of the students who are experiencing reading difficulty. ... We must remember that "below the bar" are individual children with different needs, and behind them are teachers who need policies that support thoughtful teaching and learning. (p. 235)

And once again, the children and youth who are really hurt most by all this continue to be those in poorer communities and in grossly underfunded and poorly staffed schools (Kozol, 1991; Payne & Biddle, 1999). Ironically, and sadly, the trend toward "takeovers" of "failing schools" not only appears to be ill-conceived but may end up hurting even more the very students these measures are presumably designed to protect (Malen, Croninger, Muncey, & Redmond-Jones, 2002). There is no way, I would suggest, that reasonable arguments can be made that students in such systems have equal opportunities to learn. And it may well be that this kind of argument will ultimately win the day in court; see, for example, the case of *Williams et al. v. State of California* (American Civil Liberties Union, 2001; Chapter 5, this volume).

Perhaps the saddest part of all is that there is really no solid evidence that these kinds of heavy-handed, test-based, accountability policies really work in meaningful and enduring ways. First of all, if they did, we would have improved public education long ago or at least stopped hearing about how bad our schools are. Simple logic suggests that given all the efforts since the early 1970s with minimum competency approaches, coupled with more recent efforts ostensibly focused on world class standards, we would have heard about substantial improvements by now.²

The counter argument, of course, is that even with this history of testing and accountability, there has never been a real emphasis on hooking rewards and sanctions to student test scores such as in the current high-stakes environment. But what do we know from this environ-

ment over the past half decade or more? It still doesn't seem to be working in any substantial or consistent way, and, even based on analyses by high-stakes accountability advocates is unlikely to ever work in satisfactory and tidy ways (Brady, 2003; Finn, 2002).

Emerging nearly every month are new studies documenting the failure of high-stakes testing and accountability strategies and their predictable fallout (e.g., Casas, 2003; Cimbricz, 2002; Gandara, Rumberger, Haxwell-Jolly, & Callahan, 2003; Grant, 2002; Livingston & Livingston, 2002; Mathison & Freeman, 2003; Mintrop, 2002; Rigsby & DeMulder, 2003). Yet there are some studies still trying to find the pony in the manure pile. At best, the results from these studies are mixed. For example, reporting on the June 2002 conference "Taking Account of Accountability: Assessing Politics and Policy" held at the Kennedy School of Government (Harvard University), Olsen (2002) notes the teeter-totter nature of claims, even of interpretations by competent policy analysts of data on the same long-term reform initiatives (e.g., Chicago public schools).

Other competent policy analysts who have spent considerable time studying the difficulties of major, standards-based, high-stakes accountability initiatives (e.g., Elmore & Burney, 1997) know how simplistic and wrong-headed a test-driven accountability model can be. Commenting on the No Child Left Behind Act of 2001, for example, Elmore (2002) notes:

This is an "accountability bill" that utterly fails to understand the institutional realities of accountability in states, districts, and schools. ... In the history of federal education policy, the disconnect between policy and practice has never been so evident, nor so dangerous. Ironically, the conservative Republicans who control the White House and the House of Representatives are sponsoring the single largest—and most damaging—expansion of federal power over the nation's education system. (p. 1)

My guess is that the empirical case for the effectiveness of current high-stakes accountability will suffer the same fate as research, for example, on Head Start, class size, and whether money matters in public education. Depending upon ideological alliances, the same data or similar data will be used to draw opposite conclusions. Currently, this is illustrated well by the work of Amrein and Berliner (2002), on the one hand, and that of Raymond and Hanushek (2003) on the other.

Sadly, whether pro or con, these sorts of warring, empirical studies are all based solely on point-in-time, on-demand test scores, and pretty small, incremental, average gains or losses at that. Does anyone serious-

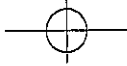
ly believe that what a kid scores on an on-demand test really represents anything more than a small sample of highly contextualized paper-and-pencil behavior, ostensibly having something to do with teaching and learning, and, these days, a lot to do with a heavy dose of test preparation? Surely, what matters more is the cumulative impact of teaching and learning and the future potentials of each child and young adult in the care of our public schools. A test score is a mighty poor indicator of a human being's potential to become all that he or she can be (Scheffler, 1985). (I will have more to say about this in my concluding chapter.)

In my view, therefore, debates about whether or not high-stakes testing and accountability "works" cannot be resolved by looking at outcomes based on high-stakes tests! The issues "at stake" are far deeper than that, and arguments based on test scores diminish considerably serious moral and pedagogical concerns. Independent of their analysis of test score results, for example, Amrein and Berliner (2002) propose an interesting social sciences version of the Heisenberg Uncertainty Principle: "The more important that any quantitative social indicator becomes in social decision-making, the more likely it will be to distort and corrupt the social process it is intended to monitor" (p. 5). And as they report in their study, as has been reported in so many others, there is much support for this hypothesis:

Because there are numerous reports of unintended consequences associated with high-stakes testing policies (increased drop-out rates, teachers' and schools' cheating on exams, teachers' defection from the profession, all predicted by the uncertainty principle), it is concluded that there is need for debate and transformation of current high-stakes testing policies. (p. 2)

Although it seems like a never-ending job description, educators must continue to hold the "accountabilists" accountable for their rationales and actions, and that is the primary purpose of this book. I first ran across the neologisms "accountabilism" and "accountabilist" in a wonderful little critique of accountability, 1970s style, by Martin, Overholt, and Urban (1976). I think these are still useful terms to signal an alluringly simplistic ideology that has captured the attention of many, both right and left, on the political spectrum. And notwithstanding claims today of new-found accountability concepts and practices, Martin, Overholt, and Urban's rendition of the ideology in the 1970s is remarkably contemporary:

The notion that all or most educational objectives should be couched in behavioral terms, the requirement that pedagogy be competence- or per-



formance-based, the insistence on a strategy of education evaluation which limits itself to that which can be observed and measured, and a call for the use of techniques of behavioral control which depend on an assumed instrumental relationship between means ... and ends. (p. 3)

Today, for example, if you go to the Fordham Foundation web site (www.edexcellence.net) and click on "standards, tests, and accountability," you'll find great optimism about what is essentially an old paradigm: "For too long, U.S. education has lacked meaningful standards and avoided real accountability. Thankfully, this is starting to change. The quest for educational accountability relies on a three-legged stool: standards, assessment, and consequences." In effect, whether we talk objectives or standards, minimum competency or world-class, the core ideology continues to hold promise for test-driven changes in teaching and learning, and rewards or punishments (high-stakes) to alter behaviors of educators and their students who apparently wouldn't do so otherwise.

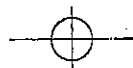
Remarkably absent in this rationale is the need for ongoing professional development so prevalent (and costly) in the corporate world. And remarkably implicit is the pernicious assumption that many schools, educators, and children can make great changes unfettered by the inequitable and often miserable conditions and circumstances within which they exist, try to teach, and try to learn.

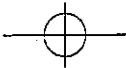
None of us who have contributed to this book are against standards. None of us believe that there are no good uses for test-based assessment or evaluation strategies. All of us believe that appraisal is important, that the public has a right to know how well schools are educating their children, and that the very nature of education itself should model good inquiry and reflective practice. All of us, however, are deeply concerned about what happens when heavy-handed accountability schemes get superimposed on the complexities of schooling practice.

In setting forward their "95 theses" on reforming evaluation practice, Lee Cronbach and his colleagues (1981) were well aware of the important distinctions between evaluation as educative and accountability as manipulative:

Accountability emphasizes looking back in order to assign praise or blame; evaluation is better used to understand events and processes for the sake of guiding future activities.

A demand for accountability is a sign of pathology in the political system. (p. 4)





Rooting out this pathology and developing more responsible practices require using what we already know about good teaching, learning, and assessment, and about the conditions and circumstances within which good educational practices can flourish, and drawing on centuries of experience with accountability of one form or another in public education.

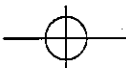
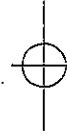
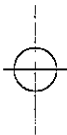
In constructing this book, therefore, I first put forward a series of eight claims for more responsible accountability and a short paragraph describing each. I then asked contributing authors to unpack these claims (Chapters 1-8, this volume). Then I received their papers, and, in several cases, modified my initial claims and descriptive paragraphs. One of my goals was to get clearer about some of the complex and contentious issues in high-stakes accountability arguments, and the contributors to this book helped me to do just that. Here, then, are the claims and the authors.

CRITICAL CLAIMS

Responsible accountability systems must pay attention to lessons of the past. This is an overarching claim that in many ways frames all of the chapters of this book. Specifically, though, the first two claims I will make target the contributions of two historians of education, and six more claims are unpacked by the six additional contributors to this book.

First, *responsible accountability systems must pay attention to the history of accountability paradigms and critical analyses of their successes and failures.* To what extent has accountability always characterized public education in one way or another, and have any of these efforts ever really been successful? I have suggested above that if accountability efforts had ever been successful in any substantial and sustained way, we should by now have established a mantra in this nation about how good schools have become instead of quite the opposite. What is needed, however, is a more nuanced and historical look at how accountability has been construed in the history of public education, and that is what Larry Cuban sets out to do in Chapter 1. Indeed, Cuban extracts interpretations from his historical analysis that problematizes, in instructive ways, the notion of whether or not accountability has "worked."

Second, *responsible accountability systems must pay attention to the history of schooling and attempts to change and standardize behavior through impositions of consequences.* Why is it so ingrained in Western thought that the only way to guarantee people will do well is to threaten them with



dire consequences for doing worse? The basic rationale for high-stakes assessments and accountability³high school graduation tests, for example⁴is that without them, teachers and students will not take better teaching and learning seriously. Without the threat of punitive consequences, it is argued, people have no incentive to do better. But, as Nancy Beadie shows us in Chapter 2, these coercive notions have a long and troubling history in this nation, and there are both moral and strategic lessons to be learned.

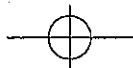
Third, *responsible accountability systems must be clear that standards and assessment are not arbitrary and are subservient to reasoned judgment about the educational aims and ideals of schooling.* Perhaps present-day assessment and high-stakes accountability systems live up to this claim. But if they do, then, by inference, the aims and ideals of schooling in America must be a watery broth of numeracy and literacy soup. Nothing resembling educationally rich visions of critically minded human beings, human beings fully capable of continued learning and intellectual growth, is suggested by narrowly defined, on-demand tests and, more important, by the equally narrow pedagogical practices focused on passing such tests. How might a reasoned argument be framed about what matters, and what ought to matter, when it comes to fully educated human beings? In Chapter 3, Harvey Siegel constructs just such an argument and shows (via one case study) how present-day assessment and accountability systems fail in light of his argument.

Fourth, *responsible accountability systems must be sensitive to the complexity of the social, political, and economic circumstances within which this nation expects its public schools to function.* The inequities that continue to mark our most troubled schools demonstrate that racism and classism are still alive and functioning in our social fabric and, thus, in today's (and likely tomorrow's) schools. Ironically, the rhetoric of "closing the achievement gap" has been picked up by both the conservative right, who may be more interested in privatizing public schools by showing continued failure to close the "gap" and by the liberal left, who believe that high-stakes testing and accountability is the only way to keep demonstrating the "gap" and, therefore, putting pressure on schools to do something about it. What truly needs closing is the "rhetorical gap," the gap between urging the closure of the "gap," and the commitment of the enormous resources it would take to ameliorate the pernicious outcomes of poverty, racism, and discrimination. Having said this, there is much room for schools to do far more than they have in dealing with these effects, even if the "rhetorical gap" is not closed. It is tricky business talking about this without being labeled negatively in some way by

various proponents from either the left or the right. How can a convincing argument be constructed that both doesn't let schools off the hook (within reason) yet doesn't put the onus on them — use them as convenient scapegoats — to solve what are still the nation's problems to solve? In Chapter 4, Pedro Noguera continues this critique and offers compelling and concrete policy initiatives that would need to be implemented if policy makers were serious about really narrowing the achievement gap.

Fifth, responsible accountability systems must be as focused on schooling conditions and equitable opportunities for student learning as on what students should be learning. Given the documented disparities in quality of education due to structural features like tracking within schools and the "savage inequalities" characterizing many of our inner city and urban (and some rural) schools, it is hard to defend the proposition that today's schools offer equal opportunities for student learning. This has legal as well as moral implications. It is morally indefensible for public schooling in a social and political democracy to stand by with full knowledge that significant numbers of children (usually poor and of color) are disenfranchised from a quality of education being received by wealthier (and usually White) students³ not only in nearby districts or in the same state or even out of state but sometimes even in the same school district or, worse yet, the same school! It may well be the case that a definitive case will need to be won in the courts before any headway on authentic equity and opportunities for learning can be accounted for in our public schools. Jeannie Oakes, Gary Blasi, and John Rogers have firsthand experiences with just such a challenge, and in Chapter 5 they share with us the substantive arguments and prospects for legal (if not moral) resolution.

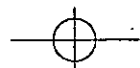
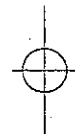
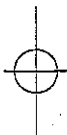
Sixth, responsible accountability systems must attend to all the core purposes of public schooling in a political and social democracy. If "what gets tested is what's taught" is the central rationale of high-stakes assessment and accountability systems, then it is astonishing and appalling that civic education is virtually absent from these systems. If teaching students their moral and intellectual responsibilities as critical and informed citizens in a democracy is a central purpose of public schooling, than surely we would have a responsible system of assessment based on civic education in all of its complexity. But we don't, and in Chapter 6, Roger Soder takes up the question "Why not?" Other critical areas could have been the focus of the larger issue here, for example, the arts (music, drama, pictorial art, etc.). Given the centrality, however, of civic education to sustaining a dynamic system of governance in a



democracy (e.g., Barber, 1993), it seemed appropriate to use this domain of education in investigating the logic of accountability and the ideological issues that cut to its core.

Seventh, *responsible accountability systems for educational practices must themselves be guided by sound educational practices*. Although there is still much to be learned, much is already known about best educational concepts and practices and the reciprocal connections between teaching, learning, and assessment. Unfortunately, both the theory and practice of traditional, on-demand, high-stakes testing and accountability systems are at odds with a great deal known about good teaching, learning, and assessment. High-stakes testing/accountability schemes result in well-known, negative fallout (e.g., teacher demoralization and student dropout, particularly economically poor students and students of color) and counterproductive educational practices (e.g., narrowing of the taught curriculum by teaching to the test, student retention in grades, and turning away from good pedagogical practices like individualization and cooperative learning). Essentially, test-driven reform models require belief in the assumptions that (1) all children are ready to be assessed at the same time in the same way on the same things, (2) useful information of various types is not already available to teachers for making good instructional decisions about individual students, and, most astonishingly, (3) education should be driven by assessment rather than the other way around. To think alternatively, as Linda Mabry shows us in Chapter 7, requires debunking "test-driven improvement" rationales, arguing for the importance of professional judgment by educators and for active involvement of students in their own evaluations, and advocating sound, "educationally driven" pedagogical practices including individualization and accumulating and using relevant information on each student over the course of their K-12 educational experiences.

Finally, *responsible accountability systems must be as focused on the continued learning of educators as they are on that for students*. The paltry amount of resources directed at the professional development for teachers and administrators in the nation's schools is disgraceful. No major cutting-edge company today could maintain its edge in the marketplace if it spent less than 5% of its resources (about what most school districts have available) on professional development, yet educators are now called upon to teach to higher and higher standards with precious little in the way of in-service training. In Chapter 8, Pat Wasley addresses these big questions: What do we know about the impact of good teaching on the quality of schooling? What are the characteristics of the kind of professional development and learning community that can make a



difference in both the quality of educators' work lives and in all the valued outcomes for students?

Understanding and internalizing the implications of these claims, I believe, are crucial to once and for all giving up an accountability paradigm has never served well the children of our nation. And until we give it up, we are not likely to get on with real progress toward more responsible ways to appraise the education of our children and the success of our public schools. So let's see how each author has helped us to further understand these eight claims.

NOTES

1. Although the evidence is circumstantial, a strong case has been made that principal Betty Robinson's suicide may well have been linked to the accountability-based threat she perceived of her school's closure and/or being fired from her position. See "Was Betty Robinson" (n.d.) and the compelling letter written by her curriculum director colleague Kate Kirby (2002).

2. I hasten to add that, contrary to what I believe is politically motivated hype, I don't think public schooling is all that bad, generally, or that it has gotten worse over all these years. See, for example, the analyses of Bracey (2002) and Berliner and Biddle (1995). However, the plight of economically poor students and students of color in our public schools, particularly in our urban areas, is still a disgrace by any standard.

REFERENCES

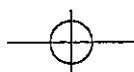
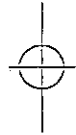
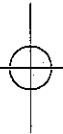
- American Civil Liberties Union. (2001). Court allows CA schools lawsuit to proceed as a class action, making case largest education lawsuit. Retrieved January 23, 2003, from <http://www.aclu.org/StudentsRights/StudentsRights.cfm?ID=9161&c=31>.
- Amrein, A.L., & Berliner, D.C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved March 28, 2002, from <http://epaa.asu.edu/epaa/v10n18>.
- Barber, B.R. (1993, November). America skips school. *Harper's*, 39-46.
- Berliner, D.C., & Biddle, B.J. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. New York: Longman.
- Bowman, D.H. (2001, September 5). Delayed again: Ariz. moves high school exit exam to 2006. *Education Week*, 21(1), 27.
- Bracey, G.W. (2002). *The war against America's public schools: Privatizing schools, commercializing education*. Boston: Allyn & Bacon.
- Brady, R.C. (2003). *Can failing schools be fixed?* Washington, D.C.: Thomas B. Fordham Foundation.

Introduction

- Buly, M.R., & Valencia, S.W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis*, 24(3), 219-239.
- Casas, M. (2003). The use of standardized tests in assessing authentic learning— A contradiction indeed. *Teachers College Record*, <http://www.tcrecord.org>. ID Number: 11211. Retrieved October 15, 2003.
- Chicago teachers boycott CASE. (2002, Fall). *FairTest Examiner*, 16(4), 5.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice. *Education Policy Analysis Archives*, 10(2). Retrieved 1/10/02 from <http://epaa.asu.edu/epaa/v10n2.html>.
- Cronbach, L.J. & Associates. (1981). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Dillon, S. (2003, May 22). States are relaxing education standards to avoid sanctions from federal law. *The New York Times*, A25.
- Elmore, R.F. (2002). Testing trap. *Harvard Magazine*, 105(1). Retrieved January 21, 2003 from <http://www.harvard-magazine.com/on-line/0902140.html>.
- Elmore, R.F., & Burney, D. (1998, May). The challenge of school variability: Improving instruction in New York City's Community district #2. *CPRE Policy Bulletin*.
- Finn, C.E. (2002). Making school reform work. *The Public Interest*, Summer #148, 85-95.
- Gandara, P., Rumbberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives*, 11(36). Retrieved October 9, 2003 from <http://epaa.asu.edu/epaa/v11n36/>.
- Gehring, J. (2002, May 8). Vote to award diplomas defies state testing policy. *Education Week*, 21(34), 3.
- Goodnough, A. (2001, June 14). Strain of fourth-grade tests drives off veteran teachers. *New York Times*, A-1
- Grant, S.G. (2002). An uncertain lever: Exploring the influence of state-level testing in New York State on teaching social studies. *Teachers College Record*, 103(3), 398-426.
- Gwinnett Citizen, The. (n.d.). *Was Betty Robinson trying to tell us something?* Retrieved 1/16/03 from <http://gwinnettcitizen.com/Editorial/bettyrobinson.html>.
- Haney, W. (2000). *The myth of the Texas miracle in education*. *Education Policy Analysis Archives*, 8(41). Retrieved December 4, 2001 from <http://epaa.asu.edu/epaa/v8n41/>.
- Heubert, J.P., & Hauser, R.M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Hoff, D.J. (2000, December 15). N.Y.C. probe levels test-cheating charges. *Education Week*, 19(16), 3.
- Jacob, B. A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23(2), 99-121.
- Jones, M.G., Jones, B.D., Hardin, B., Chapman, T.Y., & Davis, M. (1999). The



- impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81(3), 199-203
- Keller, B. (2002, January 16). Austin cheating scandal ends in no-contest plea, fine. *Education Week*, 21(18), 3.
- Kirby, K. (2002, May 5). Principal who died was gifted, devoted educator. *The Atlanta Journal and Constitution*, JJ.7.
- Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.
- Livingston, D.R., & Livingston, S.M. (2002). Failing Georgia: The case against the ban on social promotion. *Education Policy Analysis Archives*, 10(49). Retrieved 12/10/02 from <http://epaa.asu.edu/epaa/v10n49/>.
- Manzo, K. (2001, May 16). Protests over state testing widespread. *Education Week*, 20(36), 1, 26.
- Malen, B., Croninger, R., Muncey, D., & Redmond-Jones, D. (2002). Reconstituting schools: "Testing" the "theory of action." *Educational Evaluation and Policy Analysis*, 24(2), 113-132.
- Martin, D.T., Overholt, G.E., & Urban, W.J. *Accountability in American education: A critique*. Princeton, NJ: Princeton Book Company.
- Mathison, S., & Freeman, M. (2003). Constraining elementary teachers' work: Dilemmas and paradoxes created by state mandated testing. *Education Policy Analysis Archives*, 11(34). Retrieved September 25, 2003 from <http://epaa.asu.edu/epaa/v11n34/>.
- McNeil, L.M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Mintrop, H. (2002). The limits of sanctions in low-performing schools: A study of Maryland and Kentucky schools on probation. *Education Policy Analysis Archives*, 11(3). Retrieved 1/16/03 from <http://epaa.asu.edu/epaa/v10n49/>.
- Nolan, S.B., Haladyna, T.H., & Hass, N.S. (1992) Uses and abuses of achievement test scores. *Educational Measurement: Theory and Practice*, 11(), 9-15.
- Olsen, L. (2002, June 19). Accountability studies find mixed impact on achievement. *Education Week*, 21(41), 13.
- Orfield, G., & Kornhaber, M.L. (Eds.) (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. New York: The Century Foundation Press.
- Payne, K.J., & Biddle, B.J. (1999). Poor school funding, child poverty, and mathematics achievement. *Educational Researcher*, 28(6), 4-13.
- Raymond, M.E., & Hanushek, E.A. (2003, Summer). High-stakes research. *Education Next*, 48-55.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Belknap/Harvard University Press.
- Rigsby, L.C., & DeMulder, E.K. (2003). Teachers' voices interpreting standards: Compromising teachers autonomy or raising expectations and performances? *Education Policy Analysis Archives*, 11(44). Retrieved November 18, 2003 from <http://epaa.asu.edu/epaa/v11n44/>.



Introduction

- Rose, L.C., & Gallup, A.M. (2003). The 35th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 85(2), 41-52.
- Sandham, J.L. (1999, April 7). Exam-testing breaches put focus on security. *Education Week*, 18(30), 20, 23.
- Schrag, P. (2000, August). High stakes are for tomatoes. *The Atlantic Monthly*, 19-21.
- Schemo, D. J. (2002, October 15). Sidestepping of new school standards is seen. *New York Times*, A21.
- Scheffler, I. (1985). *Of human potential*. Boston: Routledge & Kegan Paul.
- Shepard, L.A. (1991). The effects of high-stakes testing on instruction. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Shepard, L.A., & Smith, M.L. (Eds.) (1989). *Flunking grades: Research and policies on retention*. London: Falmer Press.
- Sirotnik, K.A. (2002). Promoting responsible accountability in schools and education. *Phi Delta Kappan*, 83(9), 662-673.
- Walsh, M. (2002, March 6). Court declines case challenging promotion-assessment ties. *Education Week*, 21(25), 30.
- Winerip, M. (2003, May 28). The changes unwelcome, a model teacher moves on. *The New York Times*. Retrieved May 28, 2003 from <http://www.nytimes.com/2003/05/28/education>.
- Whitford, B.L., & Jones, K. (Eds.). (2000). *Accountability, assessment, and teacher commitment: Lessons from Kentucky's reform efforts*. Albany: State University of New York Press.