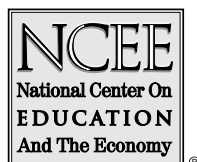


Building the Capacity of Schools, Districts and States to Educate All Students to High Standards: The Case of the America's Choice School Design

*Marc Tucker, President
National Center on Education and the Economy*

September 2, 2002

This article was commissioned by the RAND Corporation and will appear as a chapter in a book to be published by RAND on the subject of scaling up comprehensive education reform programs



Building the Capacity of Schools, Districts and States to Educate All Students to High Standards: The Case of the America's Choice School Design

Marc Tucker, President

National Center on Education and the Economy ¹

This article was commissioned by the RAND Corporation and will appear as a chapter in a book to be published by RAND on the subject of scaling up comprehensive education reform programs

Since the early 1980s, education policy in the United States has been increasingly focused on student performance. Driven by the disappearance of jobs for young people with low skills, rising demand for access to college, reports revealing the poor performance of American students in international comparisons of educational achievement and the disgust of legislators when increases in funding for the public schools failed to lead to significant improvement in student achievement, policy makers have been demanding much greater accountability for student performance. Whereas, in the past, both state and federal legislators tended to defer to the education profession and local decision makers on education issues, state and federal policy makers are now making it clear that they intend to set clear standards for student performance and they expect the profession and local boards of education to do whatever is necessary to make sure students meet them.

At both the state and federal level, policies are in place that call for raising achievement levels of those at both the top and bottom of the performance

¹ The paper is a chronicle of organizational learning. It represents the distillation of 12 years of hard won experience, not just my own, but the experience of everyone involved in the elementary and secondary education program of the National Center on Education and the Economy from its founding in 1988 to the present. I am happy to assume the responsibility for any errors of fact or interpretation to be found here, but whatever credit is due for learning from our successes and failures must be shared with all of my colleagues. To select just a few for special mention is by implication to diminish the contribution of the many I do not single out, and for that I offer my apologies, but I cannot fail to point to my colleagues Judy Coddling and Peter Hill as the two people with whom I am in almost constant exchange on the issues treated here and to Phil Daro, whose legacy with respect to the design of our curriculum and professional development strategies will be with us for a long time. Thanks, too, to Judy Coddling, Tom Corcoran and Peter Hill for their extensive and very helpful comments on this paper. Finally, I would like to thank Tom Glennan, who many years ago made it possible for me to indulge my interest in building the capacity of educational organizations and who has encouraged me in that interest ever since.

distribution. For the first time in American history, schools and the agencies that support them are being called upon to raise the top tier of student performance and to substantially narrow the distance between the top performers and the worst performers.

The penalties for failure are unprecedented. Now, for the first time, students can be denied a high school diploma if they fail to meet high standards. Younger students can be denied promotion to the next grade if they fail to perform at stipulated levels. But the consequences of failure to make real progress are no less consequential for school faculties and district central office staff.

The No Child Left Behind Act of 2001 exemplifies the public mood. Under the terms of the Act, if a school has not made satisfactory progress for five years in a row, the state must take the school over, turn it over to a private management contractor to run it or convert it to a charter school. If a school district has not made satisfactory progress for only two years in row, the state must withhold funds, institute new curricula, replace district personnel, remove schools from the district's jurisdiction, appoint a trustee to run the district, abolish the district or restructure it.

But we need to ask what capacity low-performing schools have to get their students to these new standards. What capacity do the district offices and state departments of education have to help them? Holding schools, districts and state agencies responsible for improving student performance will produce little if they do not have and cannot rapidly develop the capacity to do the job. Denying students the credentials they need to get a job and go on for further education is immoral if the institutions on which they depend cannot provide them the services they need to get to the standards. The question of capacity is the central question provoked by the standards and accountability movement. Perhaps the framers of the No Child Left Behind Act and all of its near relatives believe that educators have always known how to greatly improve student performance, but were just waiting for someone to put more pressure on them to do it.

I doubt it. I think they know that most educators have been doing their best and simply don't know how to get better results within current resource and organizational constraints. Few if any large city systems have been able to produce sustained high achievement for large numbers of children from low-income families. The same is true for rural areas serving low income children. In fact, much the same is true for individual schools. This observation holds true not just for regular public schools, but for schools supported by vouchers or charter arrangements as well.

So the most important question facing American education today is what, exactly, it will take to produce the transformation in student performance that is called for by the new policies, and who will help these agencies pull it off if they lack the capacity to do it themselves? These are the questions to which this paper is addressed.

The National Center on Education and The Economy (NCEE) was founded in 1988 to develop the tools, designs, professional development and technical assistance that schools, districts and states would need to meet the challenges described above. This paper, at one level, is a chronicle of what the organization has learned about how to do that and how organizationally we have used what we have learned to get better at our work. I argue, perhaps not surprisingly, that third-party organizations like ours will prove to be a very important tool in raising student achievement across the nation in the beginning of the 21st century. In that sense, then, this paper addresses the issue of capacity in two ways: first with respect to the capacity of schools, districts and states to respond to the challenge of raising student achievement dramatically, and, second, the capacity of external organizations like the NCEE to help them reach that goal.

In Part I of this paper, I share our analysis of the characteristics of systems used by other countries that have the capacity to produce consistently high levels of student achievement at scale. Chief among them is coherence; well functioning, highly integrated systems (especially instructional systems); and explicit designs for those systems. I then offer a theory to guide the development of such systems in the United States. I hope to show that no attempt to build capacity at scale that is directed only at the school level can succeed, that the only systems that will succeed are those that are coherent from the school through the district to the state, in much the same way that such systems are coherent in those nations with more successful education programs.

In Part II, I describe our early experience in trying to help districts and states develop the capacity to produce improved student achievement and some lessons we learned from that experience.

In Part III, I identify the kinds of tools that are needed by anyone who would build the kind of capacity in the system that is called for in Part I: highly integrated instructional systems, designs for coherent educational systems, cascading blended professional development systems, and responsive technical assistance. I argue that only sustained professional development that is tightly connected to school designs and aligned with standards-based instructional systems will alter classroom practice sufficiently to result in broad improvements in student performance in low-performing schools. Finally, I argue that approaches that use only professional development and technical assistance strategies are bound to fail, and that major new investment in the development of coherent instructional systems is the necessary, though not sufficient, condition for success.

In Part IV, I talk about the ways that these resources can be deployed by a third party provider to build the capacity of schools, districts and states so that they can meet whatever challenge comes along without needing to depend on any external organizations like ourselves in the future. The immediate aim of this capacity-building approach is to enable the district or state to take over from external providers the task of supporting schools implementing design approaches to whole-school reform. But the larger purpose is to build the capacity of the state or district to support high performance schools throughout

the district or state, with all that entails, whether or not each school thinks of itself as implementing a particular model of reform. And I show why alignment of structures and systems from the school to the state is a necessary condition of stable high performance at scale.

In Part V, I discuss the advantages of having third party organizations available to provide the assistance that schools, districts and even states need to turn our education system around. I argue not that it is impossible for schools, districts and states to do the job by themselves, but rather that one is more likely to succeed with outside assistance than by going it alone.

Lastly, in Part VI, I put down a marker, pointing out the limitations on what can be accomplished by the policies and practices advocated in this paper posed by factors that are very nearly unique to the United States, chief among them the enormous inequalities in income that have come to characterize our economy in recent years and our unwillingness as a nation to provide the resources to our schools and communities to deal effectively with the consequences of poverty for our children.

Part I: The Need for Coherent Systems to Drive Educational Performance Up

We should begin by asking whether the job can be done at all. Consider that the task the nation faces, in a nutshell, is to figure out how to help schools, districts and states educate students to a much higher level of performance without substantially increasing the cost. That sounds impossible. What makes us think they can do it?

Years ago, a small group of us set out to create an organization that could help schools, districts and states build the capacity to respond to the challenge just described. We were under no illusion about the nature of the challenge, but were basically optimistic that we could meet it for two reasons. First, there was a clear precedent for an entire sector restructuring itself to meet the challenge. That was the case of the business community meeting the challenge from abroad in the late 70s and early 80s. Second, there were other countries that were substantially outperforming the United States in education, so using them as a benchmark, we should be able to learn a lot from them that would help us meet the challenge at home.

If Business Could Do It...

In the early 1980s, American businesses had to confront the reality that manufacturers in other countries were producing higher quality goods with

much shorter time to market and at much lower cost than we could. Many firms went out of business as a result, but many others rose to the challenge. Some did so by sending their production to low cost countries, but others kept production here. They completely redesigned the way they did almost everything, and learned that they could compete and win in the new environment.

In no small measure, the resurgence of the American economy was due to the willingness of American business to reinvent itself from the ground up and to profit as it did so from the successful methods of its most aggressive competitors. We were convinced that this country could do the same thing in the field of education. Report after report showed that other nations were outperforming us in education. Since 1998, when our organization was founded, we have visited 17 nations on four continents, some of them many times, trying to understand what accounts for their superior results

Benchmarking the Best

Many of the businesses that succeeded in this time of testing began by trying to understand how their newly effective competitors overseas were doing it. They did not copy them slavishly, but they did everything they could to learn as much as possible from them. We thought that this technique, which worked so well for business, might work equally well in the field of education. Not least among the advantages of this sort of benchmarking is the ability to see what it takes to make these approaches work, not in some hothouse environment, but at scale, in this case the scale of an entire nation.

The first of our trips to other nations took place in 1989, with visits to Germany, Sweden, Denmark, Ireland, Singapore and Japan. All of these countries had education systems that were outperforming our own. What was fascinating was that, different as these countries are, they have in common a very important set of features that we thought could easily account for their superior performance.

First, they had instructional systems that could properly be called systems. The list is now familiar: Clear standards, high quality examinations designed to assess whether the standards had been met, curriculum frameworks set to the standards that specified what topics and concepts were to be taught at which grade levels, a standard required curriculum (with very few electives) typically though the ninth or tenth year of school, instructional materials that fit the curriculum frameworks, and initial teacher training designed to prepare the teachers to teach the official curriculum.

We also found something else that was no less important. In most of these countries, public policy and private commitments combined to assure that most students had access to a high minimum standard of nutrition, health care, housing and other services. If income transfer programs did not take care of these problems in the family unit, then the government provided the needed resources to the municipality or the schools to address them. Coordination

between education authorities and other units of municipal government was much closer than in the United States. Differences in resources for education among jurisdictions within the nation were smaller. The national education authority was stronger in those nations than either our state level or national level of authority, and the local (school district) level of authority was much weaker.

When we compared these common characteristics with prevailing practice in the United States, the differences were striking and illuminating. The United States had no standards for student performance, much less common standards for all students. Our teachers abhorred the idea of teaching to the test, whereas teachers in these other countries could not conceive of examinations that were not used for the purpose of assessing whether a student had mastered the official curriculum (which was, of course, what the teacher was teaching). Because there were no official curriculum frameworks in the United States, publishers of instructional materials were forced to publish books reflecting the vast variety of preferences of the members of the textbook selection committees of the states and districts they sell to, with the result that our texts treated a great range of topics at a very shallow level. In the absence of standards and curriculum frameworks, the texts had come to define the curriculum, with each teacher making his or her selection of the topics to be studied, often on the grounds of comfort with that topic. Schools of education rarely offered coherent programs to aspiring teachers, much less programs that would systematically prepare them to teach to state standards (if any such standards existed), because our traditions of academic freedom leave each professor more or less free to teach what they wish to teach.

Thus the curriculum for teacher education was unhinged from the curriculum of the schools, tests were unhinged from the curriculum, and the curriculum was unhinged from any standards the students were meant to achieve. It became patently obvious to us that neither the United States nor any of its school systems had anything that could realistically be called an instructional system, all the parts of which fit together in a harmonious whole.

But that is not all. In most of these other countries, all students had strong incentives to take tough courses and work hard in school because the standards applied to everyone and they led to real consequences for the students. Students who did not meet high academic standards were denied entrance to university. Students who did not meet vocational standards were denied access to all but the worst jobs. In the United States, access to a vast range of jobs required no more than seventh of eighth grade literacy and access to most post-secondary institutions required no more than that.

It is very important to acknowledge both the strengths and limitations of this analysis. I am arguing here that the nations of the world have been playing out a great natural experiment and that those that have chosen to use the system I have been describing have been able to produce much higher levels of achievement along the whole distribution of student performance than this nation has. That is a very strong argument for adapting the system they have

developed for use in this country, especially since nothing else we have tried has worked very well.

But we need to acknowledge that the results we get, while better than we have had, may not be as impressive as they have gotten. That is because the distribution of income in the United States is more unequal than in any other country we have studied. Whereas students in other countries typically had access to decent housing and good medical care, no student in the United States could count on that as a right. Poverty loomed much larger in this country as a problem of the students in school than in most other countries we researched. This will affect everything from the physical ability of students to read the printed page (because they need eyeglasses), to the size of the vocabulary spoken at home to the amount of time that parents are available to their children, all important determinants of academic achievement over which the school has no control.

None of these conditions are major impediments to achieving school goals as long as the aim of schooling is to sort out the student population and provide a minimum level of literacy to the bottom half of the distribution. But they became mighty problems when the aim is getting all students to high standards.

I would therefore argue that we have no choice but to emulate those nations that have compiled a superior record in educating their students, and we should expect to improve our performance by doing so, but we should not necessarily expect to meet or beat their record as long as we are a world leader among the developed nations in the field of income inequality and fail to take adequate measures to protect our children from the consequences of that situation by providing them, with adequate nutrition, health care, housing, and pre-school education.

Among the frequently voiced objections I have heard to the idea of adapting the education policies and practices of other countries to the needs of the United States is the assertion that “they are homogeneous and we are extremely diverse.” That is true. We are the most diverse large nation on earth in terms of racial, ethnic and religious background. But the second most diverse large nation on earth is Australia, a nation that scored very high on the most recent large-scale international comparative study of educational achievement, the OECD PISA study. Racial prejudice is clearly a problem in the United States, but we cannot claim that it is impossible for us to match the performance of the better performing nations because of our racial, ethnic and religious diversity.

A Small, Underdeveloped Country Trumps the Best of the Best

Perhaps the most interesting example of the phenomenon I am describing is Singapore. Americans know Singapore as the tiny island nation that puts people in jail for throwing chewing gum on the sidewalk. But students of education policy know Singapore as the nation whose education system was somehow

transformed, in less than 20 years, from that of a backwater third world developing country into one of the most successful education systems in the world. In the first administration of the Third International Mathematics and Science Study, Singapore was in the top tier of performers, beating out many highly developed countries in the competition for best performance in mathematics and science, both in math and science literacy and in advanced math and science.

Not only did they beat our Advanced Placement students, but it also turns out that the bottom 15 percent of Singaporean students performed above the median of all students tested worldwide. Their bottom 15 percent outperformed our bottom half. These youngsters come from families that have emigrated mainly from India and Malaysia and occupy the lowest social ranks in Singapore (and are often referred to as the 'sweepers'). Singapore's ability to educate these students to a world class standard is especially important in the light of what was said earlier about the unique challenges facing educators in the United States.

We visited Singapore to find out how they did it. If they are using some particularly novel pedagogical approaches, we did not find them. We saw very little that we would describe as a real pedagogical innovation. But it was all very well crafted and, most important, it all fit together. The design of the Singaporean education system was notable for its coherence. It was as if it had been consciously engineered so that all of its pieces joined together in such a way that each element reinforced the others. And each piece was very high quality. There are at least three secrets to the Singaporean success story in education. The first has to do with people — they assign their best teachers to the students who need them the most, and those teachers work with their students as many hours a week as it takes to help them succeed. The other two secrets are benchmarking and design. They are intimately related.

Less than two decades ago, a new minister of education was appointed in Singapore. Determined to radically upgrade the system, he did something characteristically Singaporean; he created a team to benchmark best education practice worldwide. The team was expected not only to identify the best practices, but also to carefully sift through those practices in order to design a system that would suit Singapore. They approached the problem as Singaporeans approach all such problems, as an exercise in systems design. They did not want effective pieces of an education system; they wanted a coherent, effective system. And they got it.

Systems, Coherence and Design

These three ideas, the idea of the system, the idea of coherence and the idea of design, are essential to my argument. They are in fact three ways of expressing the same idea. Systems, when they work properly, are by definition coherent,

and coherent systems do not just happen. They must be designed. *If we want high student performance across the board, at scale, then the watchwords must be coherence, systems and design.*

It is worth lingering for a moment over the idea of coherence. I recall spending an afternoon with my colleague Judy Coddin in Singapore's center for instructional technology in the late 90s. We went there because we had heard a lot from the people in their schools about how valuable the products of this center were. Imagine my amazement when I realized that the center was producing only videotape, and had not yet begun production of any software. And even the video was hardly revolutionary. It was, in fact, a little dull by my standards, with rather flat production values. But it was immediately clear why it was valued by school faculty. It was closely tied to the standard curriculum. Because it focused on visualizing things in the standard curriculum that were hard for students to visualize unaided, sometimes using graphics and sometimes using video of relevant phenomena, the teachers thought it added a lot of value to the curriculum and they knew just how to use it with a minimum of training.

This stands in striking contrast to most of the software produced for American classrooms. Some of it is very imaginative and engaging. But the material that is most imaginative and engaging is the least likely to fit with the texts that the teacher is using as a framework for her daily lessons and even less likely to fit into the state accountability framework. Indeed, it is almost impossible for the national software producers to make products that fit, because there are simply too many texts and too many standards. On paper, of course, their software matches all standards and all texts. This is the inevitable happy product of a key word or key phrase system for matching anything to standards and means very little. In reality, it matches nothing. So the humble videotape in Singapore turns out to be much more valuable to the Singaporean teacher than much more sophisticated and expensive technology is to the American teacher.

This is far from an isolated example. It hardly matters whether we are talking about a hastily developed software product from a little known publisher, or the work of a heavily researched and evaluated, government-financed development product, if, when it reaches real classrooms in the United States, it is used, not the way the developer intended and the controlled research design contemplated, but is cropped and trimmed, pushed and pulled, taken apart and reassembled until it fits the budget, the texts, the time, the testing regime, the accountability scheme and the training of the teacher, it will bear little resemblance to the innovation intended by its originators. The Singaporean videos worked because they fit the framework of time, schedule, curriculum, texts, ancillary materials and all the rest in Singaporean classrooms.

Real coherence goes far beyond the formalistic alignment of standards, assessment and curriculum. It is what you get when culture and all the formal elements, great and small, of the environment are in sympathy with each other. It is what happens when the way classroom discipline is actually handled reflects the moral stance of the district on values. It is what happens when the reading level of the texts in all the subjects in the curriculum is correlated with the actual

reading level of the students. It is what happens when the school makes sure the parents know what the standards are that the students are expected to meet, how their children are doing, and what they can do to provide help where their help is most needed, week by week. It is what happens when the master schedule is set up so that priority in the allocation of teacher and student time goes to the students that are furthest behind and need the most help. And it is what you get when the tests or examinations you take are designed to assess whether you learned what you were supposed to learn from the courses you took, which were in turn derived from a syllabus which is referenced to the standards that you are supposed to meet. It is a matter of making sure that every aspect of the school's functioning is organized to advance its stated purposes in a harmonious way.

This most emphatically does not entail lock step curriculum or education by state dictat. It has to do with form. There are symphonies by Mozart and symphonies by Shostakovitch. They are both recognizable as symphonies, but no one would confuse the composers when listening to them. It is form that distinguishes them. Within the form, great variation is possible. Without form, cacophony is the inevitable consequence.

No one would suggest that because both Mozart and Shostakovitch wrote symphonies of a particular form that they were 'paint-by-the-number' composers. Singaporean teachers have a great deal of latitude to decide how they are going to teach the material in their syllabus. I am here suggesting that, just as it is so in the arts, it is the structure that form provides, not license, that makes fine teaching possible on a sustained, widespread basis. It will also make it possible to produce fine instructional materials that will produce excellent results, day in and day out in classrooms, organized as the constructors of the materials envisioned when they created them. It will make it possible for students to move from school to school and district to district and not be utterly lost. Coherence of the kind I am talking about is not an innovation like a new method of teaching reading. It is a necessary condition for any complex social system like a school to work.

Based on what we have learned from the countries we have studied, we can confidently say that it is impossible to achieve high student performance across the board unless there is a highly coherent instructional system in place. But it is also important to be sure that the management and organization of the education system, right down to the schools and right up to the top, are configured so as to support the standards, assessments and curriculum on which the instructional system is based. And it is no less important to make sure that the incentives faced by both students and faculty reinforce the goals of the instructional system.

The research we have done in other countries tells us that coherence does not happen by accident. Nor is it the product of professional development, even the best professional development. Professional development does not by itself produce coherent systems of instructional materials. It does not line up incentives with goals. It does not create forms of governance and school organization that line up with the structure of sound instructional systems. No, these things are the result of conscious design. Each element of the education

system must be designed to support and reinforce the others, and all must be designed to support the goal of high student achievement for all students.

Whose Capacity — To Do What?

In other nations, responsibility for the performance of the whole system rests with the ministry of education. In the United States, no agency of government has taken on the responsibilities typically assumed by the ministry of education in other countries. It is the ministry in those countries that is responsible for setting student performance standards, developing statewide examinations matched to those standards, defining an official state curriculum for the first ten grades (including frameworks laying out what topics and concepts are to be taught at each grade level in each subject), making sure that all texts and other instructional materials are keyed to the official curriculum, defining school leaving standards that also function as university entrance standards, making sure that all teachers are trained to teach the official curriculum and providing high quality training and technical assistance to schools that are not serving their students well. It is, in short, the ministry that is the “keeper of the design.”

The State?

In every state of the United States, constitutional responsibility for elementary and secondary education rests with the state, not the school district or the school. Some states have strong state departments of education, most have weak state departments of education, when measured against the functions typically performed by ministries of education in other countries. This is no accident. The American political tradition warmly embraces a deep suspicion of government. That suspicion grows as the seat of government gets more distant from the locality and that suspicion is nowhere greater than in the arena of education. Add to that the fact that the wealthier, more powerful school districts in most states have seen it as in their interest to keep the state department of education staff small and badly paid and their formal powers weak. For all these reasons, our state departments of education have largely been restricted to getting state and federal funds to the districts and performing certain narrowly prescribed regulatory functions. Historically, no state department of education has been responsible for the full range of functions that ministries of education typically perform.

This is changing. In the last few years, with the advent of the standards movement, many state departments of education have been inching their way toward actually using the powers they were granted in the Constitution, rather than delegating them to others, but they are new at it, and they typically lack the funds, personnel and experience they need to do it well.

The District?

In our system of education, the real power has resided up to now at the district level. Just ask our school principals, who, on the whole, have less real power and discretion than principals in the rest of the developed world. It is the district that hires, promotes, fires and often assigns and promotes the school faculty. It is the district that controls the school funds, and, to a large degree, the school curriculum. But, despite their high degree of effective control over the school, districts do not function like ministries of education either. They have rarely set explicit standards for student performance, established assessment systems for measuring student performance against those standards, built curriculum frameworks to specify what topics and concepts will be taught grade by grade, assured the development of instructional materials matched to a standard curriculum, or created programs to train the teachers to teach a district curriculum. Even when they have done these things, they rarely do them in harmony. The capacity of school districts, even very large ones, to perform the functions I have just described at the level at which they are performed by the ministries of most developed nations is almost nonexistent.

The School?

Now consider the school. There are some education reformers who insist that it is part of the teachers' professional responsibility to develop their own standards, build their own assessments, create their own curriculum frameworks, write their own instructional materials and design all of the instruction that they offer their students. It is bit like expecting physicians to develop their own pharmaceuticals. We know of no nation in the world that has managed to create a successful education system that way. Teaching, when well done, is a demanding, exhausting job. The functions just described are very expensive and time consuming to perform. To expect teachers to do it all is to expect teachers to do more than is humanly possible.

The reality is that the majority of American schools are places where the pieces simply do not fit together and there is very little that the teachers in those schools can do about it. But American educators have rarely seen this as a problem. The faculties of many American schools are accustomed to a kind of controlled chaos. For them, it is unremarkable that the tests they use do not match the state standards, the textbooks do not match the tests, and each teacher is given credit for professional education that is unrelated to the needs of the school. American schools are the setting for a constant stream of improvement projects, many of which owe their existence to the enthusiasms of individual teachers, most of which can be expected to work at cross purposes with one another, few of which are expected by anyone to be elements of a concerted, multifaceted, multiyear

plan to improve performance at that school. The irony is that low performing schools are likely to be the setting for still more such special projects and programs, still more helping agents from outside the schools, still more nostrums, each of which contend with one another for the time and attention of the already beleaguered staff. In a situation of this sort, only the strongest of leaders can impose order on what has become the antithesis of a coherent program. In our experience, poor leadership and this kind of controlled chaos are the defining characteristics of low-performing schools. They go hand in hand.

It is not difficult to find schools that are more successful than the average even in the worst of districts. But they tend to be the work of renegade principals marching to their own drummer. Sooner or later, these unusually talented and driven individuals tire, and, when they do, their school inevitably reverts, in the hands of an ordinary mortal in the position of principal, to the relative chaos and unexciting performance of most of the schools around it.

The reality is that the schools, let alone districts, are in no position to perform the functions of a ministry of education for themselves. It has to be done at a higher level, and, if it is not done at a higher level, then nothing that is done at the school level will last very long.

Part II: Building Capacity to Operate a Coherent System at Scale:

The Story of the America's Choice School Design

In 1992 a new organization, New American Schools, invited groups to come forward with ideas for “break the mold” designs for schools. We decided to apply for funds in response to their announcement, because we emphatically agreed that design held the key to superior school performance, but we disagreed with their premise that schools could succeed without making significant changes in the districts and states of which they are a part. Their conception of the problem, we thought, ignored the reality that districts and states hold the key to school performance, because they establish the rules, set the culture and control the resources for the schools. We submitted a proposal to provide services to districts and states intended to enable them to field a team of senior people in the district or state department of education whom we would train to develop and support in a process of comprehensive school reform.

The emphasis in our proposal was on the word “comprehensive.” Our conception of what that word meant was summarized in a kind of litany, as follows: 1) Without explicit student performance standards, there would be no target to shoot at, and without measures of performance against those standards, the standards would be useless; but (2) if there were no change in instruction, then there would be no reason for student performance to change, so we would have to attend to the curriculum and other aspects of the learning environment;

but (3) if the management and organization of the school did not change to reflect a determined focus on student performance, then every attempt to improve performance would be defeated; but, even if all these things happened inside the school, (4) if something were not done about students who come to school hungry, who have no home and who, in other ways, lack essential supports and services; then they will not learn, and finally, (5) if we fail to engage parents and the community, then they will reject the whole effort and it will fail. We called the italicized phrases our five design tasks, because, we said, each of these arenas required a carefully thought through design and all of the designs needed to be welded together into an integrated, coherent system.

We thought that the way we would be able to help the teams assembled by districts and states develop good designs and implement them in each of these areas was to identify best-of-breed organizations in each area and have those organizations deliver the professional development and technical assistance the districts and states needed. So we reached out to organizations like the Learning Research and Development Center to partner with us on curriculum, Xerox to partner with us on leadership and management, Apple to partner with us on technology, and the Center for the Study of Social Policy to partner with us to help schools make the right connections with the social service system.

The reviewers that NAS assembled were kind enough to overlook the fact that we had written a proposal that did not conform to the RFP; we won the largest contract awarded by NAS. New York City, Pittsburgh, Rochester (NY), San Diego, Vermont, Washington State, Kentucky and other cities and states joined with us, naming teams to be trained and committing to a standards-based approach to comprehensive reform.

Focusing on School Design

But this approach failed to give us the effects on student performance that we were looking for. In retrospect, it is hard to imagine what made us believe that we could effect major changes in the schools in a large district by involving a half a dozen senior people in the central office in a program of intensive professional development. All we had to offer them were general principles. We had no instructional system, we had no school designs, and we had no blueprint for anyone to follow that would enable them to make order out of chaos. We did not know at the time that we would need all of these things and much more to succeed. But we did know that we were making very little difference on the bottom line, student performance, and we knew by then that we would have to get a lot closer to the schools to affect student performance. And, by then, mostly because of our continuing visits to other nations, we were beginning to appreciate the power of the idea of coherence in the school program.

So, when the Obey-Porter legislation establishing the Comprehensive School Reform legislation was passed, we created the America's Choice School Design.

Building on all of our experience with the program we had been running and combining it with what we had learned in our wide-ranging research on effective education systems abroad, we created designs for elementary, K-8, middle and high schools.

It might have looked to some as though we had abandoned our earlier stance and decided that the only possible path to raising student achievement on a larger scale was to go school by school, more or less ignoring the district and state levels of government and administration.

But that was not the case at all. We were still convinced that the winning approach would have to be comprehensive and systemic. And we still believed that only the states and districts can create the conditions for widespread improvement of student performance in the schools. But we had learned a lot from the work that we had done in the preceding few years.

What We Had Learned From Round One

First, we had too many partners. Each of the partners I mentioned earlier, though powerhouses in their own domains, had their own view of the world, their own methods and their own organizational needs. We were busy, it turned out, substituting one form of incoherence for another. The idea that the districts and states needed to build coherent systems of instruction, management, organization and so on was right, but we were not helping them by simply pointing them toward a set of other organizations with expertise in each of these fields, because they did not have the capacity to put it all together, with or without our help. It was as if we had given them plans for a new house and a catalogue of preferred vendors of materials and services. They needed a general contractor.

Second, we learned three related things that would turn out to focus our work in the ensuing years: First, although all of the design tasks are important, the most important components are those that comprise the instructional system. This is because the ability of the teachers to get students to standards depends absolutely on the quality of the instructional system, defined as the standards, the assessments, the curriculum framework, the instructional materials and the design of the actual classroom experience for the students, as well as the quality of the professional development they receive to help them implement the rest of the system. Second, American teachers, especially in low-performing schools, need and cannot make much progress without a fully specified instructional system, including the design, the standards, the assessments, the instructional materials and all the training and professional development needed to make the system work. It all has to be worked out in advance. Third — and this was most surprising of all — the curriculum materials that are available on the American market for this purpose cannot be used in a standards-based system. Despite all the correlations with standards being offered now by publishers of texts and other instructional materials, we discovered that they will not do the job (for

reasons that will be made clear below). New systems of instructional materials, designed expressly for the purpose, would be needed. We concluded that neither we nor our education partners would succeed unless we made a multi-year, multi-million dollar commitment to standards-based curriculum development on a very large scale.

Third, we had the wrong business model to go to scale. When we got our first contract from NAS, we reserved a substantial share of the money to share with our district and state partners, on the theory that, just as we had expenses in connection with the work, so did they, and it was only fair to share our largess with them. But this model cannot be scaled. As the numbers of schools using the model rises, the money available to pay for the schools to use it will run out, no matter how big the pot is. But if the schools and districts have to pay for the services they get, and feel that they are receiving good value for their money, there is no limit to the number of schools that can be served. So, when we created the America's Choice School Designs, we changed the business model. Henceforth, we would seek foundation and government funds for research and development and would charge our school, district and state customers the full cost of delivering the training, professional development and technical assistance needed to implement the designs. We were poised to go to scale.

There is another point to be made here. We had learned that, as with all 'soft' money that schools get, they were happy to have it as long as it lasted, but, because they assumed it would go away, they did not build any plans around it or reshape any of their operations around it. Paying for the work turned out to be a sure-fire way to ensure that our efforts would not make any difference. Almost instantly, our relationship with our customers changed. When we instituted the new policy of making our customers pay for the services they received, a few who had gotten used to us paying for everything were upset and left. Most of our former customers made the transition without complaint (some saying they had expected this years ago). New customers simply assumed they would have to pay and were not disappointed. Most important, we discovered that our hunch was right; once our customers had to pay for the services, they had to justify the cost to themselves and those who paid the bills, and, consequently, were much more likely to make the changes they had to make to fully implement the designs, because the people paying the bills wanted to see results. As soon as our customers started to pay for what they had formerly gotten free, implementation improved markedly.

The fourth problem that we had encountered in our work with the districts and states was that we were too far from the schools to have an effect on student achievement. This was actually a complex issue. First, the states and districts saw the problem as lack of capacity, all right, but not their lack of capacity. They thought the lack of capacity was in the schools, so, while they were happy to participate in the professional development we offered and accept the technical assistance, by and large, they did not think that they needed to change the way they organized and managed themselves. Second, the people we were working with, though highly placed, were not typically line managers, and it was the line managers who called the tune in the district or state. Third, they were too far

away from the schools and there were too many links in the chain between them and the faculty in the schools who could actually make a direct difference in student achievement. Lastly, the argument we could make for radical changes in the way the state or district did its work was far too weak — it was logical but abstract, hardly the stuff from which bold changes in direction come.

This last point is very important. Educators, at every level, are understandably conservative when it comes to making significant changes in the structure of the enterprise. They are justifiably suspicious of people who come to sell them ideas. They are even suspicious of educational research results. But they are not suspicious of what they can see with their own eyes.

Teachers dragged into our program by their colleagues become enthusiasts not because of what we say to them in our professional development programs but because, when they use the tools we give them in the ways that we suggest, they see their students doing things they never thought they could do. Principals become believers when their most experienced, hard bitten teachers come to them saying that they have seen performance by their students that exceeds anything they have seen in thirty years on the job. Likewise, school board members and superintendents throw their weight behind our work when they walk into classrooms and are stunned by the work that they see students doing. Data on student performance from independent, objective, third party researchers and evaluators is very, very important, but there is no substitute for seeing student work that is of a high standard, over and over again, especially from students from whom the least has always been expected.

To get this result, we realized we had to work directly with schools. There was no way around that. It is all very well, we discovered, to provide professional development to district level staff, and even to school principals, on the basis of “what the research says.” The audience, if it is well done, will be polite and even enthusiastic. But it all stops when the time comes for action, because, no matter how provocative and interesting the presentations and materials are, there is not much that the practicing educator can get his or her hands around that convinces them that it will all work when the chips are down. If this is true in general, it is doubly true of low-performing schools, which are the schools we are most concerned about. We had to demonstrate, on the ground, in the schools, beyond the shadow of a doubt, that we had what it would take for schools to make dramatic improvements in student performance. Then, and only then, would we have the legitimacy to walk into a district office or state department of education as a consultant on what it would take to go to scale district-wide or state-wide. We had to prove ourselves and we had to do it at scale, in lots of schools, in lots of places.

Capacity Must Be Built at Every Level, From School to State

Our aims had not changed at all, but our strategies for achieving them had changed greatly. In a nutshell, our original conception of what it would take to build capacity into the system to sustain widespread improvement in student

performance was decidedly anemic compared to what would really be required. None of this should have surprised us, given all we know about what was actually in place in other countries that had experienced more success.

If, then, the question is whether one gets greatest leverage at the state, district or local levels, the answer is that one can begin anywhere, but, to get strong effects at scale, the system must be coherent and aligned from the school to the state. But, it turns out that, for historical reasons, none of these levels — school, district or state — have the capacity to make that happen by itself. For us, the sensible place to begin was with designs that could be used at the school level to produce results, in the form of test scores and student work, that could be used to gain the credibility we needed to work at all three levels. We set about to do just that.

A Note on Philosophy

In a moment, we will describe what we learned about the kinds of tools that are needed to build the capacity we think necessary. But, before we do that, we need to take a step to one side, as it were, and say something about the philosophical stance with which we approach this work.

We take it as given that the goal of schooling is to educate students who not only have the basic skills — new and old — but who can be said to be deeply thoughtful, widely knowledgeable, highly skilled, committed to doing what is both right and good, and able to enjoy fully the highest achievements of human creative effort over the millennia. We want people who can think for themselves, who can both lead and be an effective member of a team, and who can cope well when confronted with the unexpected. And we hope to produce graduates who are tolerant, understand the roots of our freedoms, are prepared to take part in the political life of our nation, are willing and able to make an active contribution to their community and who have what it takes to be good parents and form strong families.

We believe that an education that has aims of this sort, that places great value on thinking and creativity and ethical and moral action, cannot be provided through instruction of a mechanical sort or by teachers who are simply doing what they have been told to do, step by step. This description of what we want for our students clearly requires a thinking curriculum and it says a lot about what we expect of our teachers.

I say this because the nature of the education one values and one's conception of what it means to be a teacher have a profound effect on the approach that one chooses to building capacity in educational institutions for raising student performance at scale. A curriculum that is truly a thinking curriculum is one that demands standards and assessments that are themselves deeply thoughtful and that reward student work which demonstrates a thorough command and deep understanding of the concepts on which each discipline is based. Building capacity among school faculties is a different task if one believes that teachers

need to be able to constantly make very important judgements about what to teach at any given moment and how to teach it than if one believes that teachers need to be taught to carefully follow a tight script as they make their way through the school day.

None of this is said invidiously. There are people for whom we have great respect who hold different views on these issues. But the views I just shared color everything we do, as well as the way we interpret our experience.

Part III: The Tools Needed for Capacity-Building

Twelve years of experience tell us that anyone who wants to develop the capacity of any level of government to steadily raise student achievement on a broad scale to internationally benchmarked standards will need to have the following tools.

Instructional Systems

The first trips that we took to Germany, Denmark, Sweden, Ireland, Japan and Singapore were organized as part of the research program of the Commission on the Skills of the American Workforce, created by the National Center on Education and the Economy in 1989. It was that research program that convinced us that coherent, standards-based instructional systems were the key to success for nations that outperformed the United States in education.

The kind of standards and examinations that attracted our attention are epitomized by the International General Secondary Curriculum Examinations offered by Cambridge University. These examinations, based on the current British standards and prescribed curriculum, are taken in about 140 nations around the globe, including Singapore. They are set to syllabi that are deeply thoughtful, and while explicit, still give teachers a lot of latitude as to where they will go deep and where they will go broad. We would wager that any teacher in the United States, given a chance to review these syllabi, would agree that they constitute a curriculum well worth teaching, and that any student who mastered such a curriculum would be justified in thinking himself or herself well educated. The final score is not based solely on the score on the end of course examination, but also on what the Cambridge people call “course work,” scores given by the teacher to work assigned to the student according to guidelines in the syllabus. These scores are moderated by the examiners at Cambridge, to make sure that all of the scores are anchored to the same scale.

The examinations are not multiple choice, machine scored tests. They require extensive writing. Because they are set to a particular syllabus, they assume that the student has read particular books, investigated particular topics and studied particular concepts and issues. After each year's examination is over, the exam is released and examples of passing papers are also released.

With such a system, students and their teachers have a very good idea of what topics and concepts need to be taught. They know what books have to be read and they know what they will be examined on. So it is possible to prepare, and effort really pays off. The standard is clear, but one cannot prep for the test in the sense of memorizing the answers to particular questions, because no one knows what those questions will be. The only way to prepare is to thoroughly master the curriculum.

No test I have seen developed in the United States approaches the quality of the Cambridge University examinations, or their counterparts in other nations we have visited.

When the work of the Commission was completed, I asked Lauren Resnick and her Learning Research and Development Center at the University of Pittsburgh to join me and NCEE to create a program to develop academic performance standards and standards-referenced examinations that could be used in the United States to produce results like those we had seen in Europe and Asia. The result was the New Standards consortium, a program of research and development on standards and testing involving our two organizations, half a dozen big cities, and more than 20 states, funded by The Pew Charitable Trusts, the John D. and Catherine T. MacArthur Foundation, the Atlantic Philanthropic Foundation and the US Department of Education, Office of Research and Educational Improvement.

The work of New Standards is still going on, but its first phase ended with the production of a set of student performance standards in English language arts, mathematics, science and applied learning for grades 4, 8 and 10; and a set of standards-referenced examinations in English language arts and mathematics, published by Harcourt Educational Measurement, for grades 4, 8 and 10. Following the first phase, New Standards created a set of grade-by-grade standards in speaking and listening and reading and writing for the primary grades. Another version of these standards was then created for English language learners. Science examinations matching the standards for middle school and high school will shortly be published by Harcourt.

These New Standards products represented a clean break with American practice with respect to both standards and testing.

Before New Standards began, virtually all standards in the United States, to the extent that there were standards, were content standards. That is, they consisted of narrative statements concerning what students should know and be able to do. One day, when we were assembling our math team, two people, one a middle school mathematics teacher and the other a professor of graduate mathematics at

a major university, stood looking at section of the NCTM standards for school mathematics. Simultaneously, each of them observed, “That is what I teach in my classroom.” Clearly, they did not teach the same mathematics. They reacted as they did because of the necessarily abstract nature of the form of statement in the content standards.

The aim of standards-based education is to get all students to a high level of performance. It is not possible to do this if neither the student nor the teacher has a clear conception of what that standard of performance is. With respect to standards, the distinct contribution of New Standards was the idea of performance standards, standards that are based on examples of actual student work that meets the standards. The New Standards performance standards consist of narrative statements of what students should know and be able to do, examples of work that meets those standards, and a commentary on each piece of student work that points out those specific aspects of the work that meet specific standards and parts of standards. The Australians had used student work to provide occasional illustrations of their standards, but New Standards took this idea one step further by creating a structure for the standards with student work at the center.

The student work in the standards was produced in response to a specific task assigned to the students who then produced the work. The standards describe the assignment and the conditions under which it was produced. The New Standards reference examinations are based, at their core, on the same sort of tasks that were developed to form the heart of the standards. Thus the standards and the examinations are seamlessly aligned, and the examinations model the kind of curriculum that is needed to enable the students to succeed on the examinations.

The examinations themselves do include short answer, multiple choice questions. But they also include a requirement to respond to complex tasks or assignments with long, written answers. All of this stands in sharp contrast to the kinds of standards and tests in wide use then and now in the states and school districts. It is still the case that most state standards are content standards, not performance standards. And it is therefore the case that most state standards are not very helpful to teachers in terms of developing curriculum or to teachers and students in providing a guide to how good is good enough.

And many states continue to use relatively inexpensive forms of fill-in-the blank tests that, far from modeling the kind of curriculum that would get students to high standards, encourages prepping for the test in ways that effectively destroy good curriculum.

Teachers elsewhere in the developed world are quite mystified when told that American teachers feel that it is unprofessional to teach to the test. They cannot conceive of tests that are unhinged from any particular curriculum. Nor can they conceive of tests of writing that do not require a student to provide an extended writing sample. In those countries, tests are examinations, of the sort described above, designed to find out the degree to which the student has mastered the

curriculum the teacher is teaching. The student provides extensive written responses to questions in the examination, and the written tests are frequently supplemented by oral examinations, often administered by the student's teacher and a teacher from another school. It is impossible to beat the test by memorizing answers. One must really know the material to do well. The closest thing to this kind of test on the American scene is the Advanced Placement test, which costs roughly ten to twenty times what the typical American test costs.

More than \$40 million was invested in the development of the initial set of New Standards products. I mention this to make the point that development of sound internationally benchmarked standards and tests is very expensive, beyond the means of almost all districts and most states. But the research that we have done in other nations shows conclusively that it is simply not possible to run a standards-based system of education successfully without high quality standards and examinations (we here adopt the definition of an examination given above: a test of whether a student has mastered a defined curriculum).

We involved 600 teachers around the nation (identified by the state departments of education as the best in their state) in the work of New Standards. The effect of this innovation on students and teachers when the teachers tried out the assessment tasks was electric. Students said to their teachers, "Oh, now I know what you want me to do. I can do that!" Teachers saw their students doing things they never thought they could do — in response to assessment tasks! As time went on, the teachers learned to post the standards in their classrooms, along with the work of their own students. Students internalized the standards and rubrics used for judging their work against the standards. They could explain in which respects their work met the standards and where it fell short, and could talk about what they needed to do to make sure that their work would meet the standards. Teachers, gathering together, began to talk about the work that their students were producing and asking each other how they were able to get their students to produce this or that. Their practice, and their efforts to improve their practice, began to revolve around the standards and the work. Suddenly, student work had become the focal point for the efforts of both student and teacher, and the effect of that was to focus teachers on their own professional practice as never before.

There was just one problem. The teachers in this group of 600 kept telling us that they could not find any curriculum materials that they could use to get their students to the standards. Lauren Resnick, co-director with me of New Standards, thought that the teachers were simply unaware of the existence of the materials that would do the job. Under this theory, our job was simply to find those materials and point the teachers in the right direction.

Others of us were not so sure. The team from the National Center on Education and the Economy decided to put it to the test, asking experts in the subject matter areas for their recommendations as to the best materials to support a curriculum intended to get students to internationally benchmarked standards like the New Standards performance standards. Nothing the experts found fit the criteria. The teachers, it turned out, were right.

How could this be? There is a bewildering variety of curriculum and instructional materials available in the United States. Surely one can find a small subset that can be assembled into a powerful standards-based curriculum that can get students from many different backgrounds to high standards!

The reason that we could not assemble a powerful standards based curriculum from available materials has to do with the nature of the American market for instructional materials. Other nations, as I mentioned earlier, have, over the years, developed curriculum frameworks to match their standards. These frameworks specify, at a minimum, what topics and concepts are to be taught at each grade level in each subject in the required curriculum.

The best of these frameworks are carefully drawn, in the following way. First, the standards are used to specify what the students must know and be able to do at the end of their basic education (not by the way, at every grade). The framers then work their way back from that endpoint to figure out what must be learned in order to produce that outcome. Each topic and concept that is necessary to achieve the standard is arrayed in the logical order in which they must be taught, so that everything that is prerequisite to something else comes before that topic or concept. Only what is essential to this set of building blocks is included. Nothing that is not necessary to the final goal is left in. Thus the curriculum is reduced to its most important components, and each of these components is addressed at length.

The curriculum that is specified in the most powerful of these frameworks is carefully balanced among skills and knowledge, concepts, and problem solving or applications.

In countries that have such frameworks, closely tied to their standards and to their assessment systems, instructional materials publishers would not dream of producing materials that did not match the framework, because they would have no customers.

In the United States, however, we have rarely had such frameworks, at any level of our education system. The result is that American publishers of educational materials have been forced to develop texts and other materials that would appeal to the widest possible audience, and the only way to do that has been for them to poll representative teachers' committees across the whole nation, asking them what they would like to see in the texts. Since there is no common framework, the answers cover an enormous range of topics, subjects, concepts, authors and so. To avoid the prospect of losing sales, the publisher must somehow get it all in there. The result is an enormous collection of the biggest, most expensive textbooks in the world. But, when you look inside, you find that each topic and concept is treated briefly, at a very superficial level.

The NSF-supported curriculum materials generally go deeper, and pay more attention to concepts, than do the more traditional materials. But it is still often hard, even with these materials, to decide what standards are being taught by

any given section of the materials and harder still to bring to the surface the relevant concepts. The Advanced Placement courses come closest in the American experience to the model we are describing, but lack some of the important features of the Cambridge examination system courses and others like it and, in any event, are offered only to a minority of American students at the end of the elementary and secondary schooling sequence.

That is what we discovered when we tried to understand what explained the observation that teachers could find no materials to get their students to the standards. Not long after we did that research, the results of the Third International Mathematics and Science Study were released. According to the authors of that study, the single most important explanation of the poor relative performance of American students in the subjects studies was the phenomenon just described, the character of the American curriculum, which the TIMSS authors characterized as “a mile wide and an inch deep.”

Thus the observation of the 600 teachers that they could not find materials that could get their students to the standards could not have been more important. The \$40 million that New Standards had spent on getting the standards and examinations right was not enough. The National Center on Education and the Economy committed to spending whatever time and money would be needed to design a curriculum for its America’s Choice School Design program and build the necessary materials in English language arts and mathematics. We did that because we knew that the schools in our network could not hope to duplicate the performance of the best performing countries unless and until our country had a matched system of standards, assessments and instructional materials that was the equal, in power, of the best in the world. If we aimed to build their capacity to get the job done, and we did, then we would have to build the curriculum materials, because no commercial publisher could afford to do so until someone had shown that the investment would pay off. We assumed that it would cost the National Center on Education and the Economy at least as much to do this as it had cost New Standards to build its performance standards and reference examinations, but we believed we had no choice.

This is not the place to describe our curriculum development program in detail, but it is important to make it clear that we are not simply talking about developing a text series in each subject. The aim is to put together a whole instructional system, a system designed to get students to the standards no matter where they start. To do that, we have had to start with the standards, then develop a curriculum framework of the kind already described, and derive a set of instructional materials from the framework, grade by grade. In order to make the materials as powerful as possible, we have built them to fit a specified set of classroom rituals and routines, generic activities that take place in an expected order as the class proceeds. As we have done this, we have designed assessments that are embedded in the curriculum that match the standards and the New Standards Reference Examinations, and we have designed professional development for teachers to match the instructional materials for the students.

A curriculum designed to take a student in at Kindergarten and brings that student up to a college-ready standard by the end of the tenth grade (that is the specification for the America's Choice Design) is going to be more demanding than 80 or 90 percent of American school curricula. So most American students are well behind where our framework says they should be when they begin our curriculum.

So we have to have a set of safety net programs — double period courses at key intervals, tutoring programs, summer school programs and so on — that are intimately tied to our curriculum, and can lift the students in America's Choice schools up from being several grade levels behind to on track in the shortest time possible. An associated assessment system is required that monitors the student at frequent intervals against the standards and signals to the teacher which students are falling behind in which areas, so the teacher can invoke the safety net features that student needs — and only those features — as soon as the student need them. We are in the process of building all of this now. It is expensive, difficult and time consuming; the research literature needed to do it faster and less expensively is full of holes, a legacy of the lack of attention that this country has paid over recent decades to the curricular needs of students who fall behind.

All these components and more need to be designed as parts of one, integrated, coherent instructional system.

These words sometimes make American teachers very uneasy because they conjure up an image of lessons that have been fully scripted by someone else in a remote office somewhere. It is the system, not the lessons that need to be fully designed. One need only look at the Cambridge syllabi and exams or the actual operation of a Japanese or Danish classroom to see almost instantly that the system assumes a high degree of professionalism on the part of the teacher and gives the teacher a great deal of discretion in deciding how to meet the needs of individual students. A syllabus — if it is the right kind of syllabus — is not a script. In the field of medicine, where no one would dream of scripting the physician's behavior or questioning her professional standing, no physician, as I pointed out before, is expected to invent his own pharmaceuticals, create his own procedures, invent her own standards of practice and so on. What the physician is expected to do is bring an enormous amount of knowledge and experience to bear on the diagnosis of the situation and needs of individual patients and, on the basis of that diagnosis, formulate a course of treatment for that patient and correct along the way to take account of developments as they unfold. We expect a teacher to perform in a perfectly analogous way. To make that work, however, we needed to formulate a set of school designs to create the structures within which the teacher can succeed.

School Designs

The school designs we created are described in detail elsewhere. We describe just enough here to give the reader a feeling for their key features.

It will not surprise the reader to learn that all the schools in the network are asked to use the New Standards Performance Standards and the New Standards Reference Examinations. When we developed the standards, few states had their own standards. Now all but one does. We explain to the schools that join our network that we expect them to use the New Standards Performance Standards as extensions of their state standards. Once our performance standards have been aligned to the state standards, they can be used to make it clear how good is good enough with respect to their content standards. Their students take both their own state examinations and the New Standards Reference Examinations. We have carefully designed the report forms for the reference examinations so that they provide feedback keyed to the standards, making them useful for diagnostic purposes.

Teachers in our program are taught how to organize their classrooms to be standards-based. With respect to the physical aspect of the classroom, this means, among many other things, hanging on the walls of the classroom the relevant standards, the rubrics for judging student work against the standards, the work of students in the class that meets the standards, and other work that is coming along. At a deeper level, it means communicating to all the expectation that they will reach the standards, helping students to internalize the standards, allocating time and other resources in relation to what is needed to get the class to standards, selecting instructional materials based on the standards and much more.

Needless to say, the English language arts curriculum and the mathematics curriculum are built around the materials that we are producing that are matched to the standards and the examinations. We have designed these courses for students who are making steady progress toward the standards, grade by grade, but we have also designed double period courses at key points in the grade sequence to enable students who are behind by several grade levels to catch up. These courses focus on the essential concepts, skills and knowledge, leaving everything that is extraneous out. Thus they are actually accelerated courses, rather than remedial courses.

All of the courses are build around a carefully conceived set of “rituals and routines” that are common to all mathematics courses and others that are common to all English language arts courses. In English, these routines include things that small groups can usefully do while the teacher is working with individuals and small groups. Once the students learn the routines in the first grade, they find out that they simply continue through the grades, so the students know what to expect when they advance through the grades, and their new teachers do not have to spend months accommodating their students to their unique way of doing things.

The routines include a standard sequence of activities in each class for each subject. The rituals include standard activities. For example, one of the activities in the mathematics program is the ‘Say Why’ ritual, in which the students are asked to say why a particular fact in mathematics is true or why a particular procedure works. This ritual both establishes an expectation that they will understand not just what to do, but why it works, as well as giving them practice in explaining the mathematics behind what they are doing.

The designs address many aspects of school organization, including the idea of having teachers follow groups of students through the grades, a system of advisement, a “house” system for breaking up large schools into smaller units so that students will be known by their teachers, approaches to developing school master schedules that facilitate the use of the safety net system.

The America’s Choice School Designs include a Planning for Results system that helps school faculty develop plans for school operations based on careful analysis of data on student performance and other key indicators. Planning for Results also includes disciplined approaches for establishing goals and targets and for searching for better ways to accomplish the goals and reach the targets. The Planning for Results system is built on an analytical structure that facilitates the development of nested targets at the student, class, grade, school and district levels.

The designs also stipulate a distributed leadership structure for the school that assures that individuals on the leadership team take personal responsibility for the key functions that are vital to improving student performance. Because these are defined roles in the school, we can design the training for the people who occupy these roles with confidence that someone in the school has been tasked with doing the work they are being trained for.

Parent and community engagement play an important part in the designs, as do student discipline and student advisement.

There are America’s Choice Designs for elementary schools, middle schools and high schools. The elementary design makes a clear distinction between the primary grades and the upper elementary grades, and includes elements in the primary curriculum and organizational design intended to make sure that all students emerge from the primary grades reading and writing well. The high school design makes a sharp distinction between the lower division years and the upper division years. Students in the lower division (Grades 9 and 10) study a standard curriculum intended to make sure that, by the time the students leave the lower division, they have the skills needed to enter college without remediation in mathematics, reading or writing. When the students have met that standard, they can proceed to an upper division program (Grades 11 and 12) that offers several different versions of ‘early college’ or can go directly to community college or technical college if they want a technical education in their upper secondary years. Thus all students in the America’s Choice system are prepared for college.

Schools are more than collections of structures, procedures and tools. They are also lively cultures, places with their own values, relationships and ways of doing things that are just as important as the formal system in shaping the quality of students' lives and the education they get. So there are parts of the designs that are intended to address these aspects of the school, too.

There is much more to the designs than I have just described, but this sketch is sufficient to convey the idea that the designs encompass not only the instructional system, but virtually every important aspect of the way a school functions.

Some components of the design are very precise. Others are better described as guidelines rather than strict requirements. We noted above our commitment to the view that teachers, to be successful, must be treated as professionals, people who do the kind of work that depends on the exercise of professional judgement of the kind that can only be exercised by the person on the spot. Nevertheless, our actual experience has been that the faculties of the schools with which we work are constantly pressing us to make the design and its requirements ever more detailed. They want us to fill in the blanks wherever possible. The more troubled the school, the more likely this is to be true.

When we ask school people what they value most about the work we are doing together, the answer is very consistent: the design has brought order and consistency to our school. They tell us that, while many of the things they are doing now are things they were doing before, they are much more productive now and make much more educational sense, because now they operate in ways that tend to reinforce each other instead of conflict with each other. Now, with this structure in place, they can see what they need to prune from their curriculum and their processes, and what is missing. They tell us that, when they are done "dropping and adding," the sum of the parts is more than the whole, whereas, before they were less. It turns out, by their testimony, that design is empowering.

In one way, that is counterintuitive, in the sense that design may constrain. But the people who use our design experience it in the same way as composers experience musical forms, as a structure that makes it possible for them to compose something much more powerful than they could put together otherwise. There is nothing more empowering than the feeling that one is effective in one's work. What the designs do is create the kinds of structures that, in other countries would be provided by ministries of education. As we noted at the outset of this paper, those structures can make schools very effective indeed, so it should not surprise us that the schools welcome the structure that these designs provide.

This, of course, begs the question of how detailed the designs should be and how insistent we should be that the designs be implemented faithfully in all cases. We will return to these questions below when we discuss the use of the designs by districts and states.

Cascading Blended Professional Development Systems, the 'Corporate University' and Responsive Technical Assistance

Clearly, our design was of no value unless the schools we worked with could be trained to implement it well.

Some years ago, when I was trying to understand how the best American firms were developing their own capacity to respond to the competitive onslaught from Japan and the other "Asian Tigers," I focused on the way David Kearns, Xerox's CEO at the time, had rescued the firm from the brink of disaster.

Kearns began by sending teams of engineers to Japan to see how the Japanese had succeeded in cutting costs, reducing cycle time and raising product quality, all at the same time. In time, Xerox perfected this process of benchmarking and made it a part of the Xerox corporate culture. But, as soon as the reports started filtering back to headquarters, Kearns put together a top team to drive the reforms within the company, consisting of a handpicked staff of senior officers and the consulting firm they had engaged. The team focused on a small core of procedures and methods that they wanted diffused through the entire organization, worldwide. The methods included benchmarking, but went far beyond that. The question on the table was how to get the message from headquarters to every nook and cranny of this far flung organization.

The answer was training. Using a system that has since come to be called "the training cascade" in business, the team that Kearns had appointed to come up with their strategy trained Kearns and the rest of the top team at headquarters. Then that team, including Kearns, trained the people who reported to them, and so on, down the line, until the lowest rungs on every ladder had been reached worldwide. This was not conceived as a static operation. Because the training included constant measurement of results against goals and standards, as well as a whole host of search behaviors to find better ways of doing things, including benchmarking, the cascade was dynamic. Each layer in the organization was empowered to find better ways of doing things within the larger framework that the new strategy created and the strategy itself was constantly under review. Each layer and business unit became a learning community.

Our aim was to create a training system that could operate at scale. We needed our own version of the training cascade.

Initially, when we had less than fifty schools in our network, all of the services were delivered to the schools by our national staff. Much of the training was delivered to the school leaders by the people who had developed our standards, assessments and instructional materials. Later, with the very rapid growth of the network, we realized that this system would no longer work. Travel costs were dominating our budget and the growing travel burden was destroying our staff.

As the professional development load grew exponentially, the time it took to do the professional development was leaving less and less time for the development staff to do their development work.

So we made two decisions. The first was to decentralize delivery of services and professional development. We established six regional offices for these twin purposes. Each office would be headed up by a regional director and include a business manager, support staff and a group of senior professional who would collectively be responsible for providing all the training and technical assistance our customers needed. These professionals typically play double roles, as cluster leaders and subject matter specialists. As cluster leaders, they have lead responsibility for eight to ten schools. As specialists, they are responsible for delivering training and technical assistance in literacy/English language arts, mathematics or leadership. In their role as cluster leader, they call on specialists in the other two areas as needed.

At the time, the number of schools in our network was doubling every year. We had no choice but to decentralize service delivery. But we were determined to maintain the quality of services we delivered. This was an enormous challenge. With the rate of growth we were experiencing, it was inevitable that most of the people delivering services would be people who had not been on our staff very long, and the decentralization of service delivery would mean that those of us in the central office would no longer be in a position to monitor the quality and consistency of our professional development and technical assistance on a daily basis. In time, each regional office would inevitably modify the original professional development curriculum to the point that it was unrecognizable and there would no longer be an “America’s Choice” way of doing things, but rather six idiosyncratic ways of doing things.

We borrowed a page from the playbook of the successful American firms of the 1980s. Faced with similar problems, they created corporate universities to make sure that their corporate strategies were reflected in the leadership training given to every layer of management in the firm. And that is what we did. We established what we called our ‘National College’ — our own corporate university — with the mission of training our regional service delivery staff nationwide, to a standard that the National College would set. When the National College was fully up and running, no member of our regional service delivery staff would be permitted to provide services solo unless that person had completed the National College program and been certified as having met our standards.

The certification program takes almost ten months to complete. The first step is a very intensive three-week “Boot Camp.” Classes start at 8:00 am and end at 9:30 in the evening. Participants study the school designs for a week and their own specialty for two weeks. Performance assessments are administered at the end of week one and week three. In the next phase, participants serve under a mentor in the field, usually a regional director, implementing a carefully designed individual professional development plan. In the last phase, the candidates present a portfolio of their accomplishments in the field and are formally

assessed in an assessment center format, primarily using performance tests. This certification program represents an enormous investment for the National Center, but we believe that our capacity to provide a high quality program depends on it.

About a year after we established the National College and began the multiyear process of implementing its design, we learned from the CPRE evaluation team that not more than half of the teachers in the America's Choice schools had changed their teaching behavior as a result of our work. The CPRE team congratulated us, saying that we had made more progress in changing teacher behavior than most external training and technical assistance providers do in the time we had had, but we were deeply disappointed and determined to improve our record.

We responded by greatly intensifying our professional development program. Our general approach, of course, is to focus the whole professional development program on the acquisition of the skills and knowledge needed to implement the design. But we have a very catholic view of that goal — ranging from exposure to the theories of Clausewitz on strategy and the views of Vygotsky on learning theory to the specific features of our Planning for Results system and the steps needed to implement our primary literacy program. Here, too, the theme is coherence. Though the professional development program is very wide-ranging, every aspect of it bears on enabling the faculty to raise student performance in the context of the plan that they have agreed to use for that purpose.

The America's Choice School Design calls for the designation of certain key people on the school leadership team to play defined roles. In addition to the principal, these include literacy coaches, math coaches, design coaches, and a person whose task is to engage parents and the community in the life of the school. Most of the professional development program is directed at the members of this team. It consists of a series of institutes and workshops, supplemented by regular supported meetings of teams of faculty members at every grade level within and across subject matter specialties, a network of principals led by the America's Choice regional director, and regular direct coaching of the principal and other members of the leadership team by the cluster leader. All of these activities are designed and scheduled throughout the year in such a way as to make sure that the professional development is continuous and logically sequenced.

Our signature is the use of the apprenticeship approach to professional development: we provide the participants with a tool, model its use, let them use it with supervision, and, critiquing their performance along the way, gradually withdraw our support until they are fully competent to use the new tool or procedure without further support. This continues for years in iterative cycles as the "apprentices" go deeper and deeper into the substance of the work, becoming ever more skilled at it.

This approach was developed as a way to raise the skill level of teachers in the schools in which we work. But, of course, though many people beyond the

members of the leadership team are directly involved in the process I just described, we cannot work directly with each teacher in this way. So, the question then arises, how does the person we have trained pass on the skills that he or she has acquired to the other members of the faculty?

Our first answer to that question was to establish a protocol for rolling out the program schoolwide. The way this works is illustrated by the pattern through which our primary literacy program unfolds. The literacy coach in the school (a full time position we require the school to establish) attends an extended institute at which he or she has a chance to understand the principles involved, observe experts modeling the expected behavior, ask questions and try some things out. When the coach returns to the school, he or she is expected to pick out a classroom that will become a model for literacy teaching in the whole school and works with the teacher in that classroom so that the teaching comes to exemplify the America's Choice approach. The coach is then expected to do the same thing, classroom by classroom, covering up to ten classrooms. If the school wants to cover more classrooms, it must designate more literacy coaches.

The model classroom becomes a place where teachers, singly and in groups can observe the new practices. Working with the coach, these teachers implement every aspect of the literacy design. Classroom by classroom and grade by grade by grade, all the literacy classrooms are eventually covered in this way and brought up to standard. In addition to working individually with the teachers responsible for literacy instruction, the literacy coach is also responsible for leading regular meetings of these teachers in which they continue their professional development and collectively reflect on and address problems of practice that are arising in their work. The literacy coaches are trained to both serve as coaches to individual teachers and to lead these collegial faculty meetings. We provide protocols for both.

An analogue of this procedure for school principals is the focused walk, a procedure that the principal can use as a guide to inspecting the whole school or a single classroom, focusing as he or she does so on a particular aspect of the design and its implementation. We provide the principal with a protocol for the focused walk, thereby making sure that it is a carefully considered, highly focused activity that both enables the principal to gather in a great deal of relevant implementation information quickly, but also to provide coaching to the faculty that is relevant, insightful and useful.

Gradually, over time, we have come to see that what is needed are a set of protocols of this sort for a number of actors that encompass a wide range of activities and that are related to one another, so that they collectively provide a structure for the members of the leadership to use as they draw together all of the separate threads of design implementation into one coherent activity.

While all of this has been going on, we have begun to explore the use of a very different — and complementary — approach to improving the depth of implementation in our schools: the use of blended professional development

systems combining face-to-face and e-learning approaches to professional learning.

Among the tools that we are currently making use of are videos of exemplary practices, simulations, games, videos of national and international experts talking to the participants, and journaling.

Given the problem we set out to solve — unacceptable levels of implementation — e-learning in conjunction with face-to-face delivery has the potential for getting directly to classroom teachers with a powerful product the quality of which does not degrade as one goes to scale.

The reader will not be surprised to learn that we are pessimistic about the value of e-learning when used alone. But, considered as one strategy along with many others, it may offer not only a way to enhance the reach of our implementation effort, but also substantial returns to scale. This is because the size of the investment made in developing the materials is unrelated to the size of the group using them, so, as the scale of endeavor increases, the unit cost for the materials decreases. We are constantly looking for investments that we can make in our own capacity with that characteristic.

Quality Control

During the first four years of the America's Choice School Design, the size of the network more than doubled each year from the preceding year. In the fifth year, the current year, we brought that growth rate down in order to put in place a series of measures designed to make sure that we could maintain the quality of service to the field when we resume rapid growth.

Quality has been an overriding concern from the day we began the program. It was the dominant motivation for creating our certification training program for our own employees, described above. If the people delivering the program to the field have not been very well trained, then nothing else we do will assure quality. The first step in any modern quality control system is to build it in from the beginning.

But we also have a whole range of means for monitoring the quality with which the design is being implemented and getting the information produced by those measures to the people who need to act on it. I mentioned earlier that we engaged the Consortium for Policy Research in Education when we started our program both to evaluate the outcomes for students and to provide feedback to us on the way the program was being implemented in the field. At several points in this paper, I described ways in which the feedback from CPRE, based on its survey instruments and fieldwork, revealed weak spots in the design or in our delivery which we moved to fix as soon as we could. The surveys include questions to school leaders and faculty about the frequency, completeness and responsiveness of the services they have received and their satisfaction with

those services. In addition, Peter Hill, our director of Research and Development, working with other members of our staff, has developed a highly detailed set of rubrics for judging the extent of implementation of the design, which has become the basis for our own regular surveys of every school in the network, which we use both to identify problems in specific schools which need attention and as the basis for correlating changes in student achievement with variations in the degree of implementation of specific features of the design. Finally, we use systematic ‘focus walks’ through the schools, sometimes led by senior staff members from our central office, to provide feedback to the schools and our staff on the quality of implementation in specific schools in the network. Some of these techniques are borrowed from the education inspection services of the countries we have researched; others are based on industrial quality control systems used by business in this country. The ‘focus walks’ are based on the work of District 2 in New York City.

Inevitably, an account of this sort of our design creates the illusion that the design is like some sort of self-winding toy that, once put down on the floor, simply executes itself and produces the desired result. That is not so. It is hard, difficult work on the ground, in real schools and communities with real people and real politics of the very local variety. We have had to give up on some schools, in the end, whose faculties were not willing to make the effort, once they understood the scope of the effort that had to be made. We have lost supportive superintendents who have been replaced with people who wanted to make their mark by introducing their own ideas, which were not compatible with ours. Our schools have sometimes had to drop out for lack of money. Sometimes teacher turnover has been so great that our efforts at professional development have simply trained teachers to work in other schools and we have been unable to train their replacements fast enough. Good principals have sometimes been followed by others who simply lacked the minimum leadership abilities to hold their school together. Some schools have taken longer to get their act together than others — and so on. The design is not self-winding. It does not execute automatically, but the record shows that when it is reasonably faithfully implemented, student performance improves markedly.

Part IV: Building Capacity at the Level of the School, District and State

Clearly, we believe in our designs. That belief is amply validated by the reports we have received from our external evaluator, the Consortium for Policy Research in Education, which show that, in the districts CPRE studied, where all or a substantial fraction of the elementary schools in the district are America’s Choice schools, students in those schools are achieving at levels significantly above the control schools. And it is validated by the very rapid growth of our

network. During the first four years of the America's Choice program, the number of schools in the network doubled every year. This last year, we held the growth down to make sure that we could put in place a series of measures intended to assure program quality. As this is written, we are just shy of five hundred schools in the network.

But that is not the end of the story. If it were the end, we would judge success simply by the numbers of schools that were implementing our design and the extent of implementation in each school. But that begs a whole series of questions. Do we believe that only schools that implement our design can be successful? Do we believe that schools that change the design over time will necessarily fail? Do we believe that schools can implement the design successfully irrespective of the actions that their district and state take and the policies they put in place? In short, do we believe that faithful implementation of the design is the one essential to success and that schools are free to implement the design if they wish?

Actually, we do not believe any of those things. In this section of this paper, we distinguish between implementation of the design as an objective and building capacity as an objective, and share our ideas about how the two might be related. We start by describing how our ideas about capacity building within the school building have evolved and then we talk about the what it might mean to work with districts and entire states to help them build their capacity — both to do the kind of work that we do and to build the capacity of the schools in their own jurisdictions.

Building Capacity at the School Level

When we began our school design work, we worked almost exclusively with individual schools. We adopted a three-year implementation cycle for elementary, middle and K-8 schools and a five-year implementation period for high schools.

It was not long before we realized that this was a mistake. At the beginning of the second implementation year for Cohort 1 schools, it was apparent that some of the schools had not made enough progress during the first year to profit from the training we had scheduled for the second year. We reconceived the progression as a progression through stages of implementation. We developed rubrics for deciding when a particular stage would be complete. The idea was to allow a school to go on to a subsequent phase only when the standards for the preceding one had been met. Thus was born, for us, the idea of standards-based implementation.

This plan, however, has proven quite difficult to implement. The schools in our network are very resistant to the idea of being 'held back' because they did not meet our standards for moving on. Only in those instances in which the

principal has left and been replaced, along with a significant part of the faculty, has it proven possible to do as we had planned.

The three year period for implementation of the elementary and middle school designs was based on the fact that that was the period funded by the Obey-Porter legislation that made Comprehensive School Reform a federal program. But it has turned out, in our experience, to be unrealistic. Now, we are telling the elementary, K-8 and middle schools that they can expect the implementation period to last five years, not three years. We have not in fact stretched out our implementation design. What has happened is that we now realize that it takes, on average, three years to get all the pieces in place, but it is often five years before the whole design is really institutionalized and the various pieces of it operating in a way that produces the full effects that can be obtained. We are speaking here of the difference between superficial implementation and deep implementation, not of the parts alone, but of all the parts together, working in true harmony, in a truly coherent system.

And we had to adjust our implementation design in other ways. We noticed, for example, that some schools were not ready to begin implementing the design at the beginning. More precisely, some high schools were not ready to begin implementation of the design as we had conceived it.

Like other organizations doing similar work, we had asked that schools interested in joining our network make sure that a comfortable majority of the faculty indicate their interest in joining our network. But it was becoming increasingly clear that this agreement did not mean much. It was usually treated in a pro forma way and few faculty members really understood what it would mean to sign up for our design. When, after we had been working with them for several months, they began to understand the degree to which implementation of the design would increase the demands on them and the extent to which it would call for them to change many aspects of their established routine, some schools, almost all of them high schools, decided they really were not on board after all.

We concluded that these schools simply did not have the capacity to begin our program at the point we had planned as the beginning point. That is tantamount to saying that they did not have the capacity to begin to build their capacity. But that was the reality. These were often very large high schools whose faculty was demoralized and in various stages of chaos, distrustful of outsiders and of each other. So we devised a program that such schools could use to engage with us that would not entail a long term commitment on the part of either party, but would enable us to work at the school in a very focused way on a problem — typically poor reading ability — that had proven intractable and underlay many of the other problems that school faced. We reasoned that, if the faculty saw that, by working with us, they could make progress on such a problem, then perhaps they would have the confidence in us to take the next step. But that time, they would also know enough about our design and the way we work to make a much more informed choice about adopting our design and committing years of

their time to working in its implementation in their school. Thus was born the idea of building the capacity of a school to go to work on its capacity problems.

And then there was the other end of the process, what happens when the school has come to the end of the implementation period and the program is over. That is how we thought about it, but it is not how the schools that joined our network thought about it. When the first cohort of elementary, K-8 and middle schools came to an end after three years, most wanted to continue. “But there is no more” we said. “But it is very important for us to remain part of your network” they said. So we worked out a package for our alumni schools that includes attendance at our national conferences (which are organized as very intensive professional development sessions attended by most of the schools in our network, occasions at which they can exchange craft knowledge, hear from experts from all over the world and learn about new developments in our design and development work), access to continuing technical assistance, access to new products not available to the general public, the opportunity to send new hires to our training programs and, so on. These schools knew that it would be vital to their sustained success to stay connected to our community, to entertain critical assessments of their operations from friends who understand what they are trying to achieve and to have access to the new products our development team was creating. They simply saw themselves as going on to a new stage in their relationship with us, not as ending their relationship because their capacity had been built. They had no such static idea of capacity and they were, of course, right. We realized that our own idea of going to scale had to include our alumni as well as those who were in the process of initially implementing the design.

What emerges from this story is the idea of capacity-building as a continuum, a journey that necessarily takes longer for some schools than others, a journey that actually never really ends. Our design now assumes that stance at its core. Like others, we advocate the idea of serious school planning based on careful analysis of student performance against explicit standards. We advocate benchmarking best practices and the widest possible search for the practices most likely to address the problems revealed by the analysis of the data. And we advocate a whole host of approaches to continuous, focused learning on the part of the faculty, entirely directed at improving the performance of the school and the students. Taken together, this is a formula for perpetual learning, perpetual capacity building and perpetual evolution.

Inevitably, if this set of procedures is carried out long enough, it will lead the faculty to adopt ideas, curriculum, techniques and approaches not contemplated in our design. And that will be just fine. In fact, it will be entirely consistent with the ethos of the design itself. When the school has developed the capacity to function at a high level, it will need less structure and prescription than it needs when its capacity is low and it is in relative chaos.

If that is so, then why are we so insistent that our schools fully implement the design in the first place? For the same reason that a seasoned composer can deviate from standard musical forms and still produce a masterpiece and the developing composer deviates from the same forms at his peril. There are many

very fine schools in this country that do not need our design or anyone else's design to do a very good job for their students. And there are a larger number of schools that absolutely require a good design to become competent organizations with the capacity to educate their students to high levels.

Here it is necessary to make a subtle point. In our opinion, all schools and society as a whole benefit from having the kind of structure that good ministries of education provide. One might think of this as The Ministry of Education Level. Here I have in mind the kinds of standards, examination systems, curriculum frameworks, instructional materials and incentive and accountability systems described above. This is the sea in which all schools swim in a country whose schools generally are performing at high levels.

Our design incorporates all that and goes a good deal further, with respect to the detail of the requirements in each of these categories. We do that because most of the schools with which we work need that level of structure, call it The Complete Design Level, in order to build their own capacity to the point where they can stand on their own and produce consistently high student performance.

This distinction between The Ministry of Education Level and The Complete Design Level, between the policy structure and the operational structure will turn out to be very important in the next section, where we talk about the idea of capacity building as it plays out in districts and states.

Building Capacity at the District and State Levels

Very early in the development of the America's Choice network, the conviction that we had long had that districts matter was completely confirmed. As early as our first year, there were districts that decided to enroll all or almost all of their schools in our network. And there were others that decided over time to enroll ever-increasing proportions of their schools in the network. And there were other places where we had only a small proportion of the schools in a district. Thus we had a chance virtually from the beginning to see what difference it makes to have a greater or lesser share of the schools in a district.

It makes a big difference. It should surprise no one that schools using our design do better when the people in the central office embrace the design and find many ways to support it in their schools.

Consider the situation when, let us say, three schools in a district of 60 schools are part of the America's Choice network. That district has its own standards and its own views about how standards should be used. The same is true of assessments and of instructional materials and techniques. Those in central office have views and policies on all these things and have a schedule of professional development offerings that they expect schools to take advantage of that are related to the preferences of the central office staff on all these points.

Of course, the America's Choice design requires implementation of a whole instructional system that is very unlikely to line up with the preferences of the central office. So any school adopting our design is immediately placed in a difficult position. Even if the central office is not hostile to the elements of our design, it is still the case that leaders of the school are in a constant state of tension with the central office on a whole host of issues arising from the conflict between the dictates of the design and the preferences of the central office. In this situation, there is very little incentive for the central office staff to be supportive of the design. They have their own agenda, and the success of schools that are released from any obligation to live with that agenda could simply undermine their authority elsewhere in the system, without redounding to their credit at all.

Now consider the situation when the top leadership of the district decides to adopt a design such as ours, not necessarily as the exclusive instructional approach to the used in the district (an issue to which I will return in a moment), but as the dominant one. When this happens, the senior people in curriculum, assessment and professional development have, we hope, been consulted. They are among the leaders in figuring out how to implement the design in their district. The success of the design is their success and they take both pride and credit for its successful implementation. Their departments are actively trying to figure out how the New Standards Performance Standards can be used to supplement and extend their standards in a useful way, how the New Standards Reference Examinations can be used to complement their assessment system, how our implementation schedule can be integrated with the district events calendar, how their professional development facilities and calendar can be made to support our design for professional development, and so on. The schools are no longer forced to choose between the dictates of our design and what the central office wants them to do. Our design is their design.

I have idealized this difference to make a point. But the difference is in fact dramatic. It is dramatic because, in the case in which the district has made a conscious choice to build a substantial portion of its program around our design, the personal stakes for the actors are greatly changed. When a substantial portion of the schools in a district are using our design, the success of the leaders of the district is tied to our success. And it is no less true that our success is tied to their success. Thus it becomes in the interest of both our team and theirs to work together to assure the success of the program, because it is both our program and theirs. It is little wonder that we get greater success in that circumstance than when the program is ours but not theirs.

But it is important to note that we are more successful in districts where we have a substantial number of schools not just because the interpersonal dynamics are more favorable, but because the result of those interpersonal dynamics is much more coherence in the instructional system and in the overall environment in which the schools work. Thus we return here to the theme with which we began. Coherence, though not sufficient, is nevertheless the sine qua non, the single most important factor, in making broad gains in student achievement possible.

And it turns out that *the way to get the coherence that really matters is to have coherence between the broad framework being used at the district level and the designs being used at the school level.*

Now, in this context, let's go back and look at the kind of coherence that one typically finds in a well run ministry of education vis à vis the kind of coherence that comes with a detailed operational design.

In the situation I just described, the district is providing policy coherence for all the actors in the district. Those schools in the district that are using the design get high levels of operational coherence. Notice that I said, "those schools that are using the design." That is not all schools, nor does it need to be. And that is the key to building capacity at the district level.

The story of our interaction with one large district will illustrate the point. The story begins with the new superintendent having just read a book that my partner, Judy Coddling, and I wrote on standards-based education. On the basis of reading the book and a visit with us, he decided that he wanted to adopt our approach as the top level strategy for his district.

He suggested, and we quickly agreed, that the best way to begin would be for us to make a full dress presentation to his board. It went very badly. In retrospect, and with much more experience behind us, it is now clear why. Large jurisdictions, big county or city education systems or states, are comfortable contracting with vendors to get specific, limited tasks accomplished that they specify, but they are very uneasy about making partnerships with outside organizations of the sort that was being suggested. To do that, as they see it, is to run the risk of being perceived as ceding responsibility and authority. The larger the entity, the less likely they are to do that.

So the superintendent suggested that we start with 14 volunteer schools (out of a total of more than 160) that would be a representative mix, some of them led by highly respected principals, who, if they became enthusiastic, would carry the message to every corner of the district. And it all unfolded much as he had predicted. Student performance rose in enough schools to attract the attention of school people throughout the district. Word spread quickly. Many of the America's Choice schools attracted a steady stream of visitors. Some of the visitors started to implement features of the America's Choice Design on their own. Soon schools from across the county were coming to the central office and the board asking to be part of the next expansion round. They included some of the top performers as well as some of the lowest performing schools. It was the system talking and not the vendor, and the people in the central office were much more interested in listening.

In the following year, close to 50 new schools in that county joined our network and began implementing the design. A close relationship developed between our senior staff and the top staff in the district, as we worked together to support the schools using the designs and to iron out a myriad of problems in the process

that, sooner or later, engaged almost every corner of this large district's central office.

Several years earlier, we had developed a district level service to match our school level service. We offered to assist districts with the development of modern accountability systems, data-based central office planning systems, decentralized resource allocation and budgeting systems, management development systems, customer satisfaction systems and many other aspects of a modern system of management based on the advances made by American business during the 1980s. The idea was to put in place in the central office all of the systems needed to support standards-based education in the schools.

The team we assembled to do this had actually done it, on the ground, in a district that became as a result a magnet for visitors from all over the Northern Hemisphere. But, in the end, we had no takers. School districts, it turned out, did not want to be told how to manage themselves. Superintendents felt that this is what they were hired to do, and their boards agree. To bring in some other outfit to coach them on how to do it is perceived as a sign of inadequacy. Again, the larger the district, the more likely this is to be true.

We learned from that. In the district whose story I was just telling, we did not offer to advise the district on any of these matters. It would have been pointless in any case, because the superintendent was as conversant with modern management practices as we, if not more so, and was making steady progress on that front. No, what we did was engage in a constant conversation with him and with the top members of his team about what was needed operationally to make the designs work in the schools that had joined our network. They brought ideas to that table and so did we. It was a genuine collaboration. And that collaboration steadily built ever more coherence into the whole system, particularly from the standpoint of instruction. In the process, here and in the other districts where we had all or a large fraction of the schools, we learned a great deal about how to organize and manage central office functions to support the process of steady district-wide improvement.

Throughout this period, we involved key people from the central office and their area administrative offices in the training we were providing to their schools. The superintendent was constantly visiting the America's Choice schools in his district and highlighting their achievements. While his support for our program produced a backlash among backers of other approaches, it also sent the clear message that there would be support from the central office for schools that chose to adopt the America's Choice design, although we felt, and the superintendent agreed, that it would be a big mistake to require schools that did not want to use the design to do so.

But the distinction between the schools implementing the design and those that were not was becoming less distinct. At the district's request, we provided leadership training early on to all the schools in the district that were not part of our network, and many features of the design were imported into those schools by their faculties. It was natural in the circumstances for the district leadership

to want to continue the work that we had begun. And that led to us training and providing office space in our regional office for cluster leaders supplied by the district. The future was clear. More and more, our mission in that district was to focus on increasing the capacity of their central office staff to support the design-based reform in *all* their schools.

But what does it mean to support design-based reform in a situation in which many schools are implementing the whole design and others are implementing it selectively? Earlier, I pointed out that it is the lowest-performing schools that are most likely to be in the most chaos and have the greatest need of structure and coherence. Another way of saying the same thing is that these are the schools that have the greatest need for a fully-specified, fully-supported school design. When laying out a capacity-building plan for a whole district, it is important to keep this in mind.

The designs we created were not designs for low-performing schools and low-performing students. To the contrary, our design standard was to create designs for schools we would be happy to send our own children to. But, as a matter of implementation design, we have always focused first on low-performing schools. This is because they are in the most need of help and because there is no message that travels through a district faster than the message that previously low-performing schools are achieving at higher levels than schools thought to be performing well.

But as close collaboration with a district proceeds, and the relationship deepens, the way the whole district operates will come more and more to reflect the principles behind the design. Moreover, the design itself will go through a metamorphosis in the district as the district makes it its own over the years. The result will be a district with policies that are coherent and also consistent with the design. The district will include among its schools some that are fully implementing a design such as ours and some that may have found their own way to success in a way that is consistent with the principles that inform the design but do not represent a faithful implementation of the design. The help we provide to a district is not intended to enable that district to turn every one of its schools into members of our network. It is intended to help them use the principles underlying our design to create a context in which all of their schools can be successful, whether or not they are faithfully implementing our design.

Clearly, what I am describing is a capacity-building process, one that applies to a large district, not a single school.

All of the considerations that I have just related came up again when we received a call from the top officials in a southern state in the late spring of 2001. The legislature had just passed a comprehensive school reform bill that required the state, among other things, to identify low-performing schools and provide substantial assistance to any school asking for it. The state recognized that it did not have the capacity to provide the kind of aid that the legislation required and

instituted a search for an organization that could provide it that led to us and our design.

The issue for both the state and for us was capacity — ours! But we figured out how we could meet both their needs and the obligations we had to the customers we already had, albeit with an enormous effort on the part of our staff. And we satisfied the representatives who came to visit us in our offices that we could and would do what they needed to have done.

Over the ensuing year, both we and the senior state staff worked hard, with good will and with great success to establish and maintain a close, collegial relationship, in the face of enormous pressures, to meet the needs of the schools and districts, respond to a deteriorating state budget picture and meet the expectations of the Governor and his staff.

From the day the relationship was initiated, the state representatives made it clear that they wanted us to build their capacity to render all necessary services on the fastest schedule possible. For that reason, they wanted us to use professional staff members of their regional education service agencies to deliver most of the services to the schools rather than our own employees. We understood and encouraged their desire to own the program and take over its management as soon as possible but we wanted to make sure that the program would be delivered to the quality standard we had set for program.

Eventually, we worked out an agreement. We would continue to work directly with the first cohort of schools in the state all the way through the multi-year, multi-phase process required to fully implement the design in those schools. We would put the regional education service agency staff delivering the program and a number of the key state officials through our certification training for providers of professional development and technical assistance. We would license the state to use our designs in their schools, provided that only certified trainers provided the service, provided the ratio between trainers and schools was the same as it is in our regular program, and providing that there was a quality assurance process in place that tracked implementation and ensured that schools were dropped from the program if they were not making an effort to implement the program.

All the way through the implementation process, we have planned our work and executed the implementation plan in very close collaboration with senior state staff.

We can look back on our year of experience with this state and see that it has illuminated the issues of going to scale and building capacity in some very interesting ways. The whole effort has been a capacity-building effort in that we have, from the start, trained central and regional state staff to do the training, professional development and technical assistance needed to implement our program on a very large scale. Regional and state officials from this state formed a substantial proportion of the participants at the “Boot Camp” I described above, the first step in our certification program. We have just received the data

from our first year's effort in this state, and it shows that the America's Choice schools are closing the gap with the schools in the state that have not been designated low-performing. The outlines of a system for building capacity for going to scale are now clear and there is no doubt that we can improve the strategy as we gain more experience.

NCEE has been working closely with governors' offices, state legislatures, the senior staff in state departments of education and the central office staff in large school districts since the day we opened our doors in 1989. Our experience over the last four years has greatly enriched our understanding of the practical issues involved in raising student performance in a standards-based, accountability-driven environment. That experience, combined with the research we have done in over 17 other nations on education system effectiveness has convinced us of the importance of working at every level of the system, from the school house to the state house, to help educators, policy makers and system managers build high quality, fully coherent standards and accountability system that include all the necessary safety nets, one of which would be the assistance program for low-performing programs, based on our America's Choice School Design.

Part V: The Case for Third Party Assistance

One could argue that the large districts and states we are working with do not need any help from us. Surely, the lessons we have learned could be appropriated by a large school district or a state directly, and that large district or state could do everything we have done to create highly coherent systems for their schools and take them to scale.

That seems very unlikely to me. I know of no state or district that could have mounted a decade long program of international research of the kind that we have mounted, or sustained year in and year out a program of instructional systems development essentially free of political pressure costing tens of millions of dollars. Fewer than a half a dozen states or cities could have assembled a staff of the kind that we have been able to assemble, from every corner of the country and abroad, to do the research, development, evaluation, development, professional development and technical assistance.

In the business world, the closest analogue to the role that we perform is the large international business consulting firms. They, too, conduct research and development on a large, international scale. They benchmark best practices worldwide, as we have been doing. They build proprietary designs, protocols, systems, training products and other tools for use by their staff and their clients, just as we do. And they invest large sums in these activities, spreading the cost across their whole client base, just as we do. Like those large consulting companies, we, too, can create an environment to attract highly capable, very experienced senior people because we can offer them the opportunity to make

change on a large scale, to be constantly learning, and to have enormous leverage at scale on an education system that badly needs what they have to offer.

We are not alone. Other organizations do the kind of work that we do. More are needed and more resources are needed both to create new organizations and to sustain and nurture those that already exist. If we are right and the need for high quality, highly coherent instructional systems and school designs is urgent, then the country would be well served if there were a number of organizations developing and offering such systems from among which schools, districts and states could choose.

Part VII: Putting Coherent Systems in Perspective

The argument of this paper can be summarized as follows. The most serious problem facing American education is the lack of capacity of our educational institutions to meet the challenge of greatly improving student achievement. The place to look for solutions is the policies and practices of nations that have been much more successful than we at educating their students to high standards and narrowing the gap between their best and worst students. When one looks at those nations, one finds standards-based systems of education the most striking feature of which is coherence — coherence with respect to the way every aspect of school functioning reinforces the others and a consistency of approach from the classroom to the statehouse, so that policies and practices reflect the same conception of means and ends and therefore reinforce each other. By contrast, the education system in the United States is characterized by relative incoherence — programs, policies and practices that, more often than not, clash and collide with one another, vitiating whatever potential they might have had individually.

This is not to say that the character of the components of the instructional system and the leadership, management and organizational systems that surround it have no bearing on the nature and extent of educational achievement of students. Some standards and some tests narrow the curriculum and leave out much of great value, for instance, and some management and organizational systems lead to more equitable outcomes than others. These are very important matters, but, in our experience, no nation has yet produced high performance at scale that does not demonstrate substantial coherence in the ways that are pictured in the graphic on the page following page 46.

Those national systems of education that appear to get the best results appear to be characterized by very thoughtfully conceived and quite coherent instructional systems that are the result of very large investments made over long periods of time. These coherent instructional systems are in turn embedded in coherent policy and operational systems that are generally consistent with the instructional systems. Fully coherent systems will have to be aligned from the statehouse to the classroom.

These points are summarized in the graphic on the next page. Clearly, the aim of both policy and practice should be to move our states and schools in them from the lower left corner of the graphic to the upper right corner. The most efficient and practical way for the states, districts and schools in this country to do this is to take advantage of the experience of other nations that have experienced greater success than we have and to create designs that capture that experience, adapted for use in this country. Third party organizations can facilitate this process in a way that is similar to the way that the established management consulting firms function in the business world, providing the tools and the training needed to enable the firm to catch up to its competitors and overtake them.

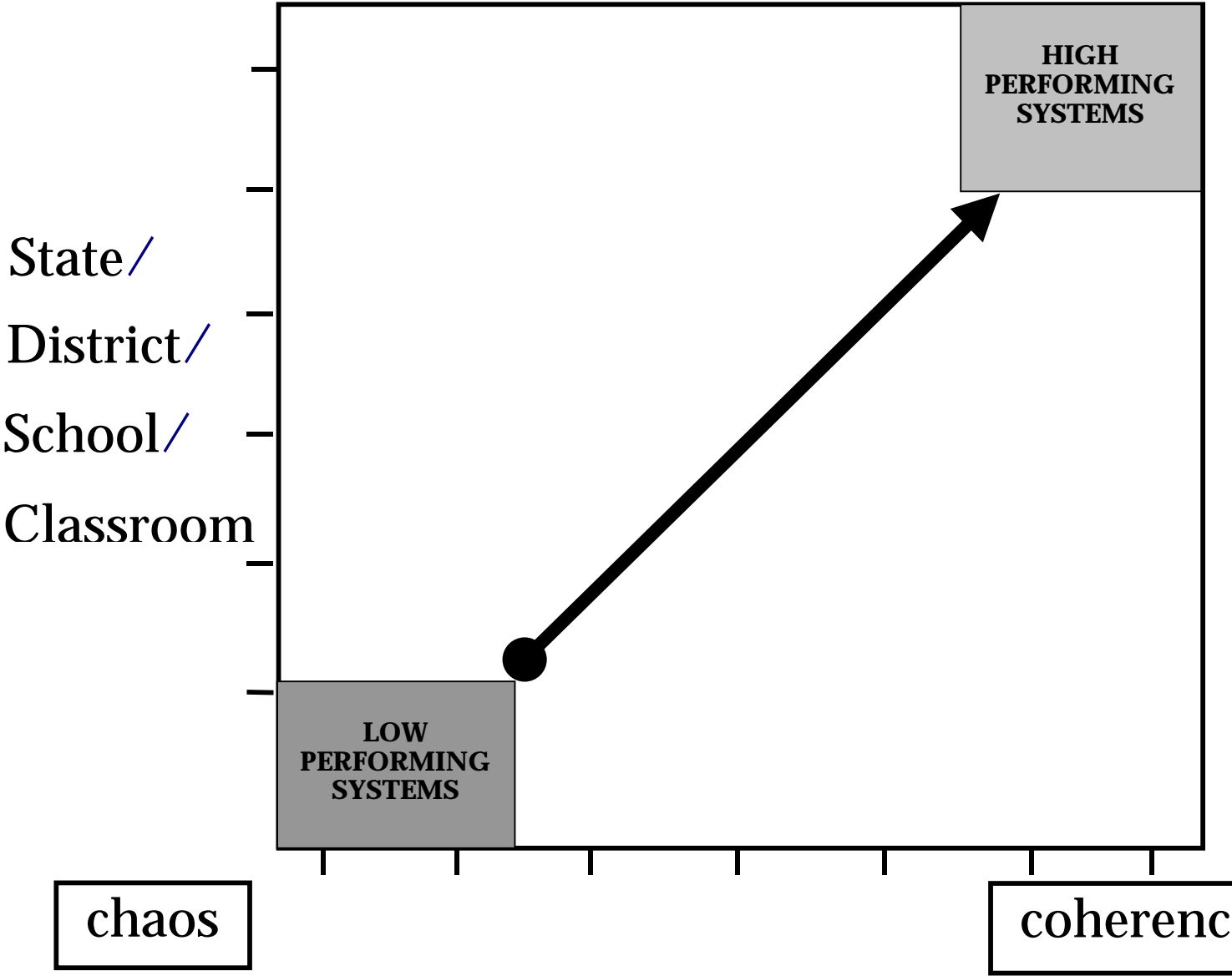
The experience gained in a lifetime of work in this field persuades me of the strength of this argument. But it is important to put it in perspective. I pointed out above that income inequalities in this country — and our unwillingness as a nation to provide the resources our schools and communities need to deal with the problems children bring to school as a result of those inequalities — will always act as a break on what the schools can achieve. We dream of what will never be if we hope for schools that can by themselves completely overcome the problems posed for children by poor nutrition, poor health, lack of stimulation in the early years, parents who use hard drugs, homelessness and abandonment.

Our willingness for many decades to write off the problems of low-performing students has left another set of legacies not faced to the same degree by the schools in other developed countries. The most glaring example is the large fraction of secondary school students who cannot read fluently. We cannot solve this problem by copying solutions developed in other developed nations, because the nations we have studied do not have those problems. We do not know very much about the solution to this problem because we have invested very little in research on it over the years. We are on our own, and we have a long way to go to develop effective, affordable solutions to many of these problems that are the result of many years of neglect.

It is also important to be clear that producing a more coherent education system will not address all the issues of capacity with which this paper began. Much has been said in this paper, for example, about what can be accomplished with sensibly conceived professional development. But there are few if any other developed nations that expect their schools to produce world class results and then flood them with teachers holding ‘emergency’ certificates who have no formal qualifications to teach their subjects. There are limits to what professional development — or any of the capacity-building strategies discussed here — can accomplish when school staff lacks the minimum qualifications to teach.

THE COHERENCE MATRIX

coherence



Standards/ Assessments/ Curriculum
Frameworks/ Instructional Materials/
Professional Development/ Planning/
Leadership/ Management/
Scheduling/ Budgeting/ Organization

More could be said in the same vein. I do not want to claim too much. I strongly believe that the capacity of our educational institutions is the biggest problem we face and building coherent systems of the kind I have described within and among those institutions is the best hope we have for raising student achievement across the board. But there are other problems we need to work on too. If we do not face them and solve them, they will continue to limit the success we can achieve by pursuing the strategies suggested here.