Approved: 2-9-2000
Date

MINUTES OF THE SENATE EDUCATION COMMITTEE.

The meeting was called to order by Chairperson Senator Barbara Lawrence at 9:00 a.m. on January 27, 2000 in Room 313-S of the Capitol.

All members were present except: Senators Bleeker, Emert and Lee excused

Committee staff present: Avis Swartzman, Revisor of Statutes
Ben Barrett, Legislative Research Department
Carolyn Rampey, Legislative Research Department
Jackie Breymeyer, Committee Secretary

Conferees appearing before the committee: Representative Glasscock
Representative Tanner

Others attending: See Attached List

The meeting of the joint House and Senate Education committee was called to order by Senator Lawrence, Chairperson. She called on Nancy Lindberg, Attorney General's office, to request a bill introduction to the House Education Committee. The legislation was introduced.

Chairperson Lawrence stated the purpose of the joint meeting was to complete the cycle of new reform plans for education. She welcomed Representatives Glasscock and Tanner to present their plan.

Representative Glasscock led off the testimony and referred to the blue book that had been distributed. (Attachment 1) He stated that each of the tabs would delineate one part of the plan which covers the four topics of, basic research, early childhood education, teacher performance and school choice research.

Representative Glasscock commented that when he and Representative Tanner approached education last summer and fall, two things were thought to be important; the one to propose reform that was significant and had meaning and reform that was backed by solid research. Research has proven that getting a good start in life makes a material difference for children in the long run. One of the prime correlations to student's learning is teacher performance and excellence. These are two elements that every parent can realize. Good teaching results in greater student learning.

Representative Glasscock took the committee through the basic skills and early childhood components of the plan and stated that Representative Tanner would speak on teacher preparation. He said that the school choice part of the plan is the only plan that is not backed by research, because there is no definitive research on school choice. There is a voucher debate every year without knowing if there is any positive correlation between school choice and student learning. The question must be answered if there is a positive, negative or neutral correlation of school choice in children's learning. A research-based study will be asked for that will create a level playing field so that at the end of the study there will be no excuses. Those that advocate for vouchers will understand whatever the results, and will accept the results. Conversely, those who have opposed vouchers, would have to accept the results should there be a positive correlation between school choice and student learning.

Representative Tanner referred the committees to the tab dealing with teacher preparation and performance and stated that he would speak of bill introduction. Several bills will be introduced in this component of the plan. He began with teacher competency which would be established by regulation of the State Board of Education. If it is determined by a school system that a newly graduated teacher from one of the teacher training institutions has a deficiency in background or training, that teacher could return to the school and receive appropriate instruction without further fees. Another bill introduction would deal with certification examination. There could possibly be another test which would determine the general knowledge capability of a teacher. Each successful person in achieving national board certification will receive a $1,000 bonus or a $1,000 commitment from the state in addition to salary. Legislation will be introduced with regard to school building report cards. Curriculum standards will be established in each of the core areas of social studies, mathematics, reading, writing and science. More

core areas could be established.

A bill will be introduced that will provide for a state requirement that each local school district develop a mentoring program especially for first-year teachers or beyond. The state will supply a stipend to assist local boards in achieving this essential program. Another area spoken of was issuance of renewal of certification. Legislation will be introduced calling for a KBI background check for all people who are entering the teaching profession. Issues that would be brought to light would be felony convictions, drug charges or sexual predator charges. There will be bill introduction with regard to a teacher service scholarship program. This would relate to length of service for a teacher, as well as dealing with hard to fill teaching positions. Legislation would state those hard to fill disciplines, meaning an under served area or a low enrollment area, would be given some assistance from the state to assist in providing an improved salary structure for those teachers.

Representatives Glasscock and Tanner stood for questions and after questions and comments, the meeting was adjourned.

SENATE EDUCATION COMMITTEE GUEST LIST

DATE: January 27, 2000

NAME	REPRESENTING
Julene Miller	AG's office
Marshal Boor	Self
Val DeFever	State Bd
Jacque Dakes	SDE
Mark Callin	KASB
Jeanyce Schmidt	House Majority Leader's Office
Kelly Levi	House Majority Leader's Office

KENT GLASSCOCK

MAJORITY LEADER
State Capitol, Room 381-W
Topeka, Kansas 66612-1504
(785) 296-7662
kentglasscock@house.state.ks.us

STATE REPRESENTATIVE
62nd District
P.O. Box 37
Manhattan, Kansas 66505
(785) 776-5353, Ext. 108

HOUSE OF REPRESENTATIVES

YEARS OF PROMISE

Kansas' Keys to Success
Glasscock-Tanner Education Plan

Overview

Preparing Kansas children for successful futures is one of the most important missions of the state. Well educated children mature into successful adults and help power our state's economy. Strong education creates a better future for all Kansans.

How we as a state ensure our children's quality education is another question. It takes a partnership between parents and the state so parents can access the best educational opportunities for their child. Ask any parent and they will tell you every child learns and grows differently – each child is unique. Our educational system should support that individuality so each child reaches his or her full potential.

This 6-year phased in plan provides parents with educational options, and focuses resources on the early years in a child's life. Research shows ages three to ten are the years of promise, the ages when children acquire their building blocks for their futures. The plan's goal is to help parents make these critical years the years of success. It concentrates reform on the early years – and ensures children have top-quality instruction. Apart from the family, nothing makes a bigger impact on young lives than great teaching, and high-quality preschool and primary education.

This plan yields great rewards for children's educational success – and higher high school graduation rates, lower crime, teen pregnancy and welfare reliance. By focusing our dollars on early childhood and primary elementary school, Kansas can save in the long run on the high costs of juvenile justice programs, prisons, welfare and many social services. This is in addition to the fact that more successful students find better jobs, and earn more money.

This is not the typical "throw money at education" reform plan. It focuses our resources, demands results and accountability, and changes students' lives. In the next century, Kansas cannot afford to sit back and accept the status quo. We will need a more highly educated and skilled workforce than ever before if we want to continue economic viability in the expanding global economy. The world of work is changing, and we have a responsibility to help parents make sure their children are ready for it. It will take four major keys:

- **Back to the basics of reading, writing and arithmetic:** Children must master the basic skills by the end of the primary grades. Right now, Kansas is not preparing children well enough to bridge into the intermediary elementary grades. The intermediary test scores show this clearly. Only 34 percent of Kansas fourth graders met the national reading assessment performance standard in 1998. In 1997 on Kansas assessments, less than half of fifth graders wrote proficiently, and only 52 percent of fourth graders were proficient in math. Research shows that if children fall behind in 3rd and 4th grades, they rarely catch up.

- **A strong start in preschool and Kindergarten:** Studies show children – especially those in poverty – make long term gains when they spend a year in half-day preschool before entering Kindergarten. Their success increases even more with full-day Kindergarten. Kansas parents of at-risk 4-year-olds need the option for half-day early learning at school and for full-day Kindergarten the year after. Parents of every child should have the choice to send their child to an all-day learning program in Kindergarten.

- **School and Teacher Accountability to Parents:** Nothing helps students learn more than great teachers – and Kansas needs to support teachers in rising to this level of performance. Since one bad teacher can cause a student to lose a year of learning and permanently lose enthusiasm for school, teachers and schools must be accountable to parents for the quality of education they provide to children. And, higher education institutions which train our state's teachers must be accountable for the quality of teaching each of their graduates provides.

- **Courage to Test New Options:** The question of whether market forces and school choice benefit children's education is far from settled. However, there's no question many students are not reaching their potential in today's education system and it is time to try some new methods to help children succeed. Kansas needs to try a research-based school choice experiment for at-risk children to find out if school choice works.

SUMMARY

YEARS OF PROMISE

Kansas' Keys to Success

- **Back to Basic Skills and Ending Social Promotion:** The Kansas State Board of Education will set standards in the basic skills areas and schools will develop plans and certification processes to ensure 95 percent of children master reading, writing and arithmetic. If these are not met by the end of the 3rd grade, students will not be socially promoted. The KSBE will also provide training in basic skills teaching methods for primary grade teachers and successful schools will receive a financial incentive.

- **Early Childhood Education:** Kansas elementary schools will offer parents the option of a half-day early learning program for at-risk 4-year-olds and all-day Kindergarten to all children.

- **Teacher Performance – Accountability in the Classroom:** The education system will be accountable to parents by:
    - requiring background checks for new teachers
    - issuing report cards to parents on the performance of their school
    - providing on-site mentors for first-year teachers
    - telling parents how their child's teacher is licensed
    - providing peer review of teachers
    - provide college scholarships in fields where teachers are needed in order to eliminate the need for out-of-discipline teaching
    - providing financial incentives for teachers to be nationally certified

  Universities and colleges must:
    - prepare teachers with technological skills
    - test each prospective teacher's subject matter knowledge to obtain teaching license
    - provide supplemental training for any new graduate who does not teach adequately
    - publicize their rate of placing graduates in teaching positions
    - publicize the percentage of their graduates who earn licenses

- **School Choice: A Research Experiment with At-Risk Students:** The Kansas Department of Education will design, implement, evaluate and assess a 4-year, research-based school choice experiment for at-risk Kansas children. The program will provide opportunity scholarships to public and private schools and answer the question of whether there is a positive, negative or neutral correlation between vouchers and improved student learning.

BACK TO BASICS

READING, WRITING AND ARITHMETIC BY THE END OF 3RD GRADE

Research Summary

From preschool through third grade, children learn to read. Beginning in fourth grade, students must read to learn. Without this key to success, children stumble with the other basic skills of writing and math. Studies show successful students must read by the time they reach fourth grade, or they start down a path of falling further and further behind. These students drop out of school and often drop into crime and welfare.

The break between the primary grades and intermediate grades – between third and fourth grades – is a divide many students are not crossing successfully. According to the Carnegie Corporation of New York, "today's fourth graders are not sufficiently proficient in reading, writing and mathematics to be able to cope successfully in the information-based, globalized economy of the next century." This shows up in the test scores. In 1992, the Governing Board of the National Assessment of Educational Progress defined achievement levels for reading and math, and only one-quarter of the nation's fourth graders scored proficiently. Kansas fourth graders took this reading test in 1998, and 34% achieved the performance standard. On Kansas standards, only 48 percent of fifth graders wrote proficiently and 52 percent of fourth graders performed math proficiently.

To ensure each student is ready in reading, writing and math for the intermediate grades, studies show elementary schools must work with each under-performing child from Kindergarten through the end of third grade. The Carnegie Corporation recommends immediate interventions with additional time and varied instruction for any child who falls behind. In addition, school districts should monitor elementary schools, and states should monitor districts to ensure that all children can read, write and do math proficiently. In the Carnegie Corporation's eyes, this should be a joint project between districts and the state. They should work together to provide professional development to better prepare teachers for this task – particularly in the area of reading.

Glasscock-Tanner "Years of Promise -- Kansas' Keys to Success" Education Plan

- By 2006, the Kansas State Board of Education must ensure that 95 percent of the state's primary students master the basic skills of reading, writing and math before entering 4th grade. The State Board will set basic skills standards in the three core areas for the end of the 3rd grade by FY 2001. Local schools will develop outcomes to meet these standards and a process to document and certify that students are proficient in each core area. These outcomes and the certification process must be filed with the State Board of Education. If the goal is not met, the State Board must create a statewide improvement plan to better educate students in basic skills.

- As a part of the accreditation process, and with the help of the $8 million federal Reading Excellence Act Grant, schools must provide summer school and after-school learning for K-3rd grade children who fall behind grade level competency in reading, writing or math. In order to give children the best chance of success, these interventions should be provided at each grade level so children are ready for the upcoming grade.

- Schools need to provide as many opportunities as possible to ensure children reach appropriate grade level proficiency in all three basic skills. If these interventions are not successful, schools may not socially promote a child who has not mastered the basic skills appropriate for his grade.

- To aid school success, the State Department of Education will provide a professional development sequence for primary grade teachers to learn the best practices for teaching basic skills. In addition, the Reading Excellence Act Grant earmarks $4 million for underperforming schools to better prepare teachers in reading.

- The success of this program should see results in the state-wide intermediary grade assessments of math in 4th grade and reading in 5th grade. The state will reward schools which successfully reach the Standard of Excellence on *both* assessments with $10 for each of the school's students. The funds will go directly to the school for use at their discretion.

BACK TO BASICS
RESEARCH HIGHLIGHTS

"The end of the third grade is a particularly important time in schooling because at that point the expectation changes from "learning to read" to "reading to learn." Although every child is unique, each having his own rate of development, interests, and learning style, with very few exceptions all children can learn to read."

"Instead of [automatically] retaining children or socially promoting them...individualized strategies should be exhausted before retention...."

Susan B. Neuman, Carol Copple, and Sue Bredekamp
Learning to Read and Write: Developmentally Appropriate Practices for Young Children
National Association for the Education of Young Children

"Elementary schools and districts need to monitor continually each child's progress toward the fourth-grade standards, beginning in kindergarten and the first grade, and intervene with additional time and varied instruction as soon as a child falls behind...."

"Children who attend an elementary school that sets high learning standards and does whatever it takes to see that children meet those standards have a better chance of leaving fourth grade proficient in reading, writing, mathematics, and science."

"States and school districts should invest adequate money, time and support in professional development of school staff. Professional development should be closely related to the school's overall strategy for meeting high standards of achievement..."

Years of Promise: A Comprehensive Learning Strategy for America's Children
Carnegie Corporation of New York

"Disruption of any of these components [in learning to read] can throw off a child's development...and could lead to difficulties that ultimately will reduce the chances that the child will finish high school, get a job, or become an informed citizen."

"The majority of reading problems faced by today's adolescents and adults could have been avoided or resolved in the early years of childhood...from birth through the third grade."

"Reforms Needed to Improve Children's Reading Skills"
The National Academies

"The poor first-grade reader almost invariably continues to be a poor reader.... The best solution to the problem of reading failure is to allocate resources for early identification and prevention.... Indeed, in the majority of cases, there is no systematic identification until third grade, by which time successful remediation is more difficult and more costly.

Joseph K. Torgesen
"Catch Them Before They Fall: Identification and Assessment to Prevent Reading Failure in Young Children"
American Educator

LEARNING TO READ AND WRITE

Developmentally Appropriate Practices for Young Children

Susan B. Neuman, Carol Copple, and Sue Bredekamp

A 1999 NAEYC Comprehensive Membership Benefit

National Association for the Education of Young Children—Washington, D.C.

Frequently Asked Questions

Since the IRA and the NAEYC released their joint position statement on early literacy, teachers and others have raised a number of questions. Here are responses to some of those most frequently asked.

Is it appropriate to set the goal that every child read by the end of third grade?

Absolutely, and the IRA/NAEYC position statement supports this goal. The end of third grade is a particularly important time in schooling because at that point the expectation changes from "learning to read" to "reading to learn." Although every child is unique, each having his own rate of development, interests, and learning style, with very few exceptions all children can learn to read. However, some children need more systematic instruction, repetition, time, and practice to learn reading skills. Some need the assistance of one-on-one attention and small-group interactions. Therefore, while our goal does not change, our methods of instruction vary to meet the needs of individual children. Teachers must expect all children to master the same ambitious content and skills, while recognizing that some children progress by different methods and at different rates.

What about children who are not reading at grade level by the end of third grade? Shouldn't they be held back?

Many state boards of education now require that children pass standardized tests in reading and other subjects, usually in third grade but sometimes even earlier, before they can be promoted to the next grade. Such policies generally are intended to make the educational system more accountable for children's learning, which is a laudable goal. Unfortunately, there are two problems with this approach. One is that standardized testing may be quite inaccurate in measuring young children's progress in reading, and too much is at stake to rely on such an index. The second problem is that research tells us that retaining children in grade is counterproductive; children rarely catch up and more often fall farther behind (Shepard & Smith 1989). Nor does the answer lie in social promotion, the practice of keeping children with their age group even though they have not learned the requisite skills and content.

The IRA/NAEYC position statement calls for teacher accountability in the form of regular, ongoing assessment of children's performance during real reading tasks. It also calls for using individualized instructional strategies (as described in the previous answer) instead of retaining children or socially promoting them. All individualized strategies should be exhausted before retention is even considered. Retaining in grade should be the last resort, not the first.

My state board of education has adopted standards or benchmarks for children's learning linked to grade levels. I'm not sure they're developmentally appropriate. What should I do?

The standards movement, which swept the country in the 1990s, has had a profound effect on education. It has the positive effect of helping teachers, parents, and children know what the learning goals and

YEARS OF PROMISE

A Comprehensive Learning Strategy for America's Children

CARNEGIE CORPORATION OF NEW YORK

Executive Summary

The years from three to ten are a crucial age span in a young person's life, when a firm foundation is laid for healthy development and lifelong learning. During these seven years, children make great leaps in cognition, language acquisition, and reasoning, corresponding with dramatic neurological changes. They develop greater facility in intellectual problem solving and abstract thinking. Their store of knowledge swells, their attention span stretches, their capacity for reflection increases. They become more proficient in their oral and written communication and better able to relate ideas and feelings to their peers. They also develop greater capability to regulate their own behavior and resolve conflict peacefully. For most children in this age period, it is not too late to overcome earlier difficulties; nor is it too early to prepare for the challenges of early adolescence and middle school.

For most children, the long-term success of their learning and development depends to a great extent on what happens to them during these years of promise. Children fortunate enough to attend a high-quality preschool or child care program and who enter the primary grades with adequate preparation have a better chance of achieving to high levels than those who do not. Children who attend an elementary school that sets high learning standards and does whatever it takes to see that children meet those standards have a better chance of leaving fourth grade proficient in reading, writing, mathematics, and science. Children whose parents create a home environment that encourages learning and who remain involved in their children's education throughout the years from three to ten earn higher grades than those whose parents are uninvolved. Children from communities that provide parents supportive programs aimed at enhancing children's healthy development and achievement and that offer out-of-school opportunities emphasizing learning do better academically than those who have not had such opportunities.

The Pattern of Underachievement

All children are born ready and willing to learn. But as they progress to and through the primary grades, a great many lose their natural curiosity and enthusiasm for learning. Millions of children are not achieving as much or as well as they could, in school or out. Most preschool programs do not prepare children for the more rigorous academic curricula that are being adopted in the primary grades. The vast majority of early care and education programs fail to meet standards of quality. As many as one-third of American children today are entering kindergarten already needing additional support to keep up with their peers. Once in school, young students are not coming close to mastering the concepts, knowledge, and skills they will need to succeed later in life.

The pattern of underachievement is especially stark for children of low-income families and children of diverse cultural, linguistic, and racial backgrounds, who by and large are not receiving the teaching and support they should have as they move from home to school to neighborhood and other settings. For them, the deck can be unfairly stacked against academic success, and the years of promise can fade to hopelessness and resignation.

Underachievement is a General Problem. But make no mistake about it: underachievement is not a crisis of certain groups: it is not limited to the poor; it is not a problem afflicting *other* people's children. Many middle- and upper-income children are also falling behind intellectually. Indeed, by the fourth grade, the performance of *most* children in the United States is below what it should be for the nation and is certainly below the achievement levels of children in competing countries. According to standards set by the National Assessment for Educational Progress (NAEP), today's fourth graders are not sufficiently proficient in reading, writing, and mathematics to be able to cope successfully in the information-based, globalized economy of the next century.

- In the 1994 NAEP assessment, nearly three-quarters of the nation's fourth graders could not meet the criteria for proficiency in reading set for their grade. Forty-two percent were unable to reach even the basic level of performance, which requires only literal comprehension of reading passages.

- In 1994, two-thirds of fourth graders could not meet the standards set for persuasive writing, narrative writing, and informative writing. On persuasive writing, nine out of ten could not meet the proficiency standards.

- In mathematics, 82 percent of fourth graders could not meet the standards on the 1992 NAEP assessment; 39 percent could not solve easy problems, such as "divide 108 by 9."

- In case studies comparing the performance of U.S. urban schools with that of Asian urban schools, the average mathematics score of fifth-grade children in only one American school was as high as that of fifth-grade children in the lowest-performing Asian school.

Has American Education Deteriorated?

Contrary to popular belief, today's schoolchildren are performing about as well as their parents and teachers did twenty-five years ago. Most American schools are managing to hold the line academically, despite the tough challenges of higher child poverty rates, frayed communities and families, and a continual stream of immigrants. Some groups — notably African Americans — are doing better than ever before. But the United States of the twenty-first century will require a much more highly educated and skilled population than it has now if it is to maintain future prosperity and ensure democratic renewal. No longer can the American education system allow so many young people to fall short of their academic promise.

Today, Americans are seeing the drastic shortcomings of an education system that is geared to the academic success of some but not all. They worry that the nation could slide into economic insecurity if their children are ill-equipped to meet the complex demands of the twenty-first century. Some may even conclude that the problems are just too big, too costly, and too overwhelming to counteract or reverse.

As confidence in the nation's education system has slipped, there has been a tendency among parents, educators, business leaders, and others to engage in mutual blaming. Such disillusionment and cynicism are mistaken. Since the 1970s, researchers have documented the many practices within families and communities as well as preschools and schools that have been shown to foster learning among children of diverse backgrounds. Today, hundreds of early learning programs, schools, school districts, teacher groups, researchers, and technical assistance organizations are demonstrating success in preventing or reversing the pattern of underachievement among children, even under the most difficult conditions. No one has all the answers yet. But enough is now known about learning and development in children between the ages of three and ten to begin making significant progress in improving the education of *every* child. What needs to happen now is to put this knowledge and wisdom to work, within and across the sectors, on a large-enough scale to make significant improvement in children's educational achievement nationwide.

Every Child Can Learn

One of the myths that has undermined school reform efforts — and damaged millions of children — is the belief that differences in the educational performance of schools are primarily the result of differences in students' inherent ability to learn. This belief is wrong. Schools fail for other reasons. Most significantly, they fail because of the low expectations they hold out for many students; the heavy reliance that schools place on outmoded or ineffective curricula and teaching methods; poorly prepared or insufficiently supported teachers; weak home/school linkages; the lack of adequate accountability systems; and ineffective allocation of resources by schools and school systems.

Circumstances of birth do indeed raise the odds against children's educational success, but these odds are not insuperable. Studies show repeatedly that children's academic performance is determined more by the time and effort they devote to learning, and by the time and effort that schools invest in teaching them, than by their inborn abilities. With the right combination of challenge and support from parents, educators, and the community, virtually every child, by the end of the fourth grade, can be reading, writing, and doing math and science at levels now achieved by only a few.

The Circle of Responsibility

The first requirement in preventing widespread school failure and underachievement is for the key learning institutions in children's lives to alter the basic assumptions about the quality of work that children can be expected to produce, so that each child is challenged to meet high expectations for learning and achievement and is given the necessary support to succeed.

Schools by themselves, however, cannot accomplish these goals for children. Schools have the primary responsibility for children's formal education, but students' educational success is influenced by far more than what happens to them in the formal system. Families and communities, preschools, after-schools, and the media all have a profound impact on children's learning, and not just during the school years — well before they enter the classroom. When a single child fails to achieve, all of these institutions are likely to be at fault. All of these institutions, therefore, have a shared responsibility to contribute positively to children's learning and development. All must begin to ask what they can do to help reverse the pattern of underachievement and bring our education system into line with our national need for a wholly educated population.

Principles of Effective Practice. Within each of these spheres of influence, there are certain principles of effective practice that have already been put to work — in parent education programs, preschools, schools, community organizations, and other key learning institutions — and that are producing positive results for diverse groups of children. From studies and evaluations of these programs, it is possible to derive certain principles of best practice that are common to all. The task force calls on all the institutions that contribute fundamentally to children's learning to start today to align their policies and day-to-day practices more closely with these common principles of effective practice, outlined below:

- Ensure, from the start, that children are ready to learn, physically and emotionally.
- Set high expectations for every child, monitor the child's progress continually, and intervene quickly when problems arise.
- Create high-quality, varied learning environments that support each child's learning.
- Provide high-level professional development to those responsible for children's education and development.

- Embed children's learning in caring and collaborative relationships with educators, parents, and other adults.
- Actively engage parents in their children's education at home and in schools.
- Accept responsibility and accountability for each child's learning and healthy development.
- Make efficient, equitable use of resources for children's education.
- Collaborate more closely with other institutions and programs that affect children's learning.

Taken together, these principles of best practice provide a broad framework for a comprehensive learning strategy proposed by the task force. If this framework is accepted by the nation, if these principles are applied within all the core learning institutions in children's lives, and if these practices are coordinated to provide children a more coherent learning experience, then all children will achieve to levels that exceed current expectations of their performance. Even if institutions do not link their efforts, there is much that each can do independently to contribute to children's educational success; the failure of one to do its job effectively, therefore, is no justification for the others to falter in their own efforts on children's behalf.

Task Force Recommendations

The task force recommendations can be encompassed within a five-point program, as follows:

- Promote Children's Learning in Families and Communities: Families are the wellspring of learning for children. To assist parents and other caregivers in fulfilling their role as children's first teachers, the task force recommends that states and communities make available to every interested family with preschool or primary grade children effective parent education and family support programs that promote learning and healthy child development. Early care and education programs and elementary schools should involve parents in their services to children. Communities should expand and improve their out-of-school programs, so that their activities are linked to children learning curricula in school. More efforts should be made to accommodate children from low-income families, children with disabilities, and children whose first language is not English. Quality standards for all community programs for children should be established and enforced.

- Expand High-Quality Early Learning Opportunities: During the preschool years, children make the developmental leaps that form the basis of later achievement. To get all children ready for school and for an education that meets high standards of achievement, the task force recommends that the nation make a commitment to expanded high-quality public and private early care and education programs for children ages three to five, supported by national, state, and local mechanisms that are coordinated to assure adequate financing.

In this mixed system of private and publicly supported programs, higher standards should be developed for facilities, staff qualifications, and overall program performance.

- Create Effective Elementary Schools and School Systems: High-quality preschools will not, however, produce lasting benefits for children if they are followed by poor elementary school experiences. The task force, therefore, recommends that states play a leading role in developing and adopting high-quality standards that specify what each elementary school student should know and be able to do across all subject areas. They should set rigorous performance standards in math, reading, writing, and science for the end of the fourth grade.

  Educators should apply the same standards of academic performance to virtually all students and use every available method to ensure that each student succeeds in meeting the requirements. Language-minority children should be offered an equal opportunity to learn the same challenging content and high-level skills expected of students proficient in English. For the small proportion of children who may not be able to meet all of the standards due to severe disabilities that affect learning, individual education plans should set reasonable goals toward meeting the highest standards possible.

  States and school districts should invest adequate money, time, and support in professional development of school staff. Professional development should be closely related to the school's overall strategy for meeting high standards of achievement and should encompass the use of effective instructional practices in the classroom.

  Elementary schools and districts need to monitor continually each child's progress toward the fourth-grade standards, beginning in kindergarten and the first grade, and intervene with additional time and varied instruction as soon as a child falls behind. School districts should monitor schools, and states should monitor districts, to provide additional support and intervention when children are not progressing toward the goals.

- Promote High-Quality Children's Television and Access to Other Electronic Media: Television and emerging interactive technologies offer a powerful, underutilized opportunity to motivate children and help them meet the higher learning standards. The task force recommends that the President, Congress, media executives, and business leaders vigorously enforce the provisions of the Children's Television Act of 1990, to ensure that every community has a variety of choices for high-quality children's educational programming throughout the week. Communities should engage local businesses as partners in efforts to create broad access to the new information technologies and

sophisticated computer applications, so that no child is denied full opportunity to use these creative learning tools.

- **Link the Key Learning Institutions into a Comprehensive, Coordinated Education System:** The discontinuities in the educational experiences of young children call for the creation of comprehensive, continuous services that link families, early care and education, and schools so that children's learning and development are reinforced from every side. State and local leadership councils or committees should create strategic plans to address the learning and developmental needs of children, based on the recommendations of this report.

Making Rational Use of Resources

Almost all of the task force recommendations can be carried out by realigning priorities and making far better use of existing monetary and nonmonetary resources — eliminating programs that do not significantly improve teaching and learning and putting existing funds toward programs that work. More public financing, however, will be needed to vastly improve the quality and availability of early care and education programs, so that children of three, four, and five receive adequate preparation for school and academic life and progress toward meeting the new learning standards. Finally, efforts must be made to reduce the dramatic disparities in public school funding across states and districts.

Many actions are needed at different levels to reverse the pattern of underachievement among the nation's children. But what is required above all is the conviction that dramatic improvement in children's learning is possible if Americans work together to build the sturdy institutions needed to assure achievement, opportunity, and coherence in the educational experience of all children. Between the ages of three and ten, children make great leaps in their intellectual prowess, social skills, and ability to manage the emotional ups and downs that are part of everyday life. If all of us could see their mental agility as easily as we observe their growing physical agility, then more Americans would believe that all children can learn to levels that far surpass our expectations.

It is within the nation's power to accomplish these results for children. If we fail to keep the promise — if we continue to focus on the most fortunate youngsters and leave the rest behind — the costs to our society in human distress, lost productivity, crime, and welfare, and in the fraying of our nation's democratic ideals, will be unbearable. The choice is ours.

For too many children, a referral to special education is a one-way ticket: too few are given opportunities to return full time to regular classrooms. In some big-city school systems, the rate of "decertifying" children from special education may be 5 percent or less.[21] In these districts, students in resource rooms and self-contained settings appear to be making very little headway in mastering basic academic skills.

Underachievement Is a General Problem

Children from all backgrounds face obstacles to successful learning — native speakers of English as well as second-language learners; children in regular classrooms as well as pupils in special education; children from middle-income families as well as those who live in poverty; children in the suburbs as well as those in the cities. If their approaches to learning are at odds with the approaches to teaching that characterize most classrooms, and their strengths and needs go unnoticed, they are at special risk of having educational experiences that are at best unsatisfactory, and at worst deeply scarring.

And indeed, by the time they reach fourth grade, the great majority of today's children have not met the standards for proficiency in reading, writing, and mathematics that have been set in this country. The United States has yet to reach professional or popular consensus on acceptable levels for student achievement, but in 1992 the Governing Board of the National Assessment of Educational Progress (NAEP) made an effort to define what the nation's students should achieve in reading and math. On the 1994 assessment, NAEP found that nearly three-quarters of the nation's fourth graders could not meet the reading criteria set for their grade.

To meet the "proficient" standard on the NAEP examinations, children must show that they can read, understand, and draw conclusions from a variety of texts selected for their age and grade. But a large percentage of students — 42 percent — were unable to reach even the basic achievement level, which requires only literal comprehension of reading passages.

Mathematics achievement is even lower. The vast majority of fourth graders — 82 percent — could not meet the standard for proficiency in mathematics on the 1992 assessment, the most recent test for which data are available. This test included problems considered to be "challenging" for fourth graders. For example, one question was, "By how much would 217 be increased if the digit 1 were replaced with the digit 5?" A large percentage of fourth graders — 39 percent — were unable to reach even the lowest achievement level, which required them to solve "easy" problems, such as "divide 108 by 9."

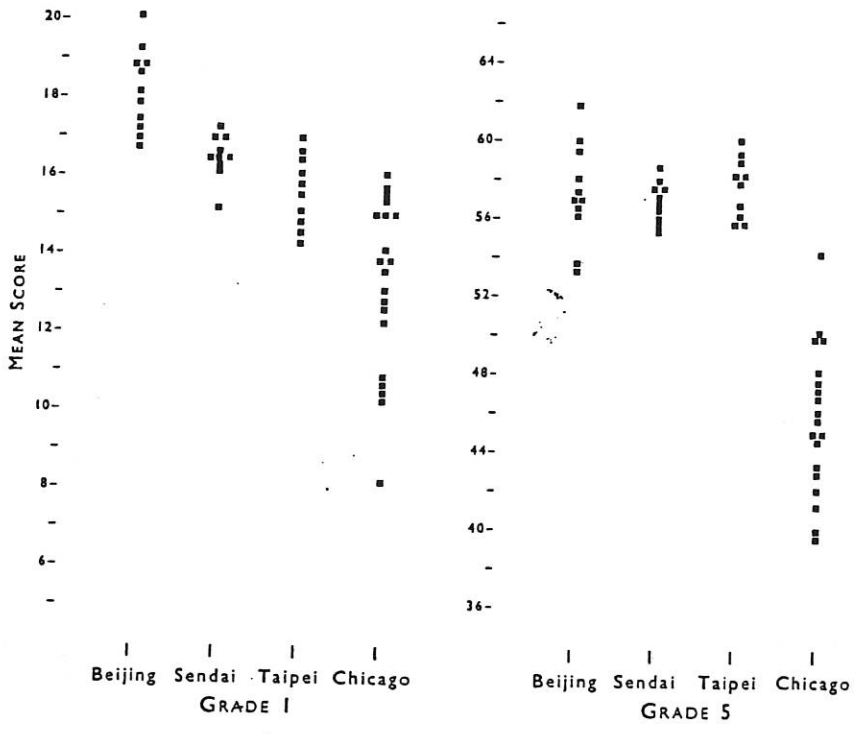
The NAEP examinations also measure writing achievement by asking students to write in ways appropriate to different purposes: *persuasive writing*, such as writing a letter to the school principal taking a stand on whether the school year should be longer; *narrative writing*, such as writing a story about an adventure as a space traveler on another

planet; or *informative writing*, such as describing a typical lunchtime at school. In 1994, two-thirds of fourth graders were unable to provide a "developed" response to any of these tasks. Nine out of ten could not meet the standard for persuasive writing.[22]

Perhaps the most telling evidence of American students' widespread underachievement emerges from a series of case studies of education carried out over more than a decade comparing schools in Minneapolis and Chicago with schools in cities

Figure 1.1

MEAN MATHEMATICS SCORES FOR FIRST AND FIFTH GRADE STUDENTS: AN INTERNATIONAL COMPARISON



The findings of a 1987 study of fifty-one schools in three East Asian cities and the Chicago metropolitan area are sobering. In the first grade, the mean for the highest-scoring American school is lower than any of the Chinese schools and is at about the median for schools in Taiwan and Japan. By the fifth grade, only one American school scored as high as the lowest-scoring East Asian schools.

Source: adapted from Harold W. Stevenson and James W. Stigler. 1992. *The Learning Gap: Why Our Schools Are Failing and What We Can Learn from Japanese and Chinese Education*, New York: Touchstone. p. 35.

1-18

Peril and Promise: Preventable Risks and Missed Opportunities

Beginning about age three, children require much less constant care than infants and toddlers, and they are also, as a group, relatively healthy. And yet, in the years of promise, American children face a range of serious risks and missed opportunities — almost all of them preventable.

Health and Safety

- One in six children under the age of five has no health insurance.
- Accidental injuries are the leading cause of death among children from three to ten — all the more tragic because most of these injuries can be avoided with safety measures such as seat-belts in cars, bicycle helmets, and "child-proof" households.
- Asthma, the most serious chronic disease of childhood, has increased by one-third since 1981. In this time, childhood deaths from asthma have doubled. These increases are related to allergens and pollutants inside and outside the home, as well as to inadequate medical care and lack of education about the disease.
- Youngsters from three to ten are at the highest risk for experiencing child abuse.

Early Care and Education

- During 1995, fewer than half of all three-to-five-year-olds with family incomes of $40,000 or less were enrolled in preschool, compared with 82 percent of children from families with incomes of more than $75,000 per year.
- Fewer than one-half of eligible low-income three-and four-year-olds receive Head Start services.
- No more than one in six three-to-five-year-olds of all income levels attends a child care center that can be considered "high quality."

in Japan, Taiwan, and China.[23] According to these and other studies, American children do not differ from Asian children in their underlying aptitude for mathematics, but their performance falls steadily behind the other groups' performance over time.[24] By the fifth grade, in the sample studied, only one American school's average score was as high as that of the lowest-performing Asian school. A few individual American students do as well as the top-performing Asian students, but they are found less often in the later grades.

NEWS
THE NATIONAL ACADEMIES
National Academy of Sciences National Academy of Engineering
Institute of Medicine National Research Council

Date: March 18, 1998
Contacts: Dan Quinn, Media Relations Officer
Dumi Ndlovu, Media Relations Assistant
(202) 334-2138; e-mail <news@nas.edu>

EMBARGOED: NOT FOR PUBLIC RELEASE BEFORE 10:30 A.M. EST WEDNESDAY, MARCH 18

This page is maintained by the National Academies Office of News and Public Information

Craig Hicks
Internet editor

Jennifer Cooke
online producer

Contact us by e-mail
news@nas.edu

Reforms Needed to Improve Children's Reading Skills

WASHINGTON -- Widespread reforms are needed to ensure that all children are equipped with the skills and instruction they need to learn to read, according to a new report from a committee of the National Research Council. An ongoing debate over which teaching method is best has diverted attention from the most important factors affecting how a child learns to read. Children need language-rich preschool opportunities, and teachers need better preparation and support to be able to guide students through the complex mix of skills that go into learning to read, the report says.

"We know what factors help prevent reading difficulties," said committee chair Catherine Snow, Henry Lee Shattuck Professor of Education at Harvard Graduate School of Education, Cambridge, Mass. "We need the will to ensure that every child has access to excellent preschool environments and well-prepared teachers. Because reading is such a complex and multifaceted activity, no single method is the answer. It is time for educators, parents, and everyone else concerned with children's education to make sure that children have all the experiences that research has shown to support reading development."

The majority of reading problems faced by today's adolescents and adults could have been avoided or resolved in the early years of childhood, says the report, ***Preventing Reading Difficulties in Young Children.*** The committee outlined critical components of a child's education from birth through third grade. For example:

> Children must arrive in first grade with a strong basis in language and cognitive skills, and be motivated to learn to read in order to benefit from classroom instruction. Preschool children need high-quality language and literacy environments in their homes and in out-of-home settings.

> Kindergarten should focus on understanding that words have letters and that letters relate to sounds; the recognition of letters; knowledge of writing concepts; and familiarity with the basic purposes and mechanisms of reading and writing. It should be designed to stimulate verbal

interaction and build vocabulary.

> First-graders should be taught to identify words using their letter-sound relationships. To achieve fluency they should practice reading familiar text, sometimes aloud. Those who have started to read independently, typically at second grade and above, should be encouraged to sound out and identify unfamiliar words.

> Beginning in the earliest grades, instruction should promote reading comprehension by helping children develop a rich vocabulary and the knowledge to use it. Curricula should include explicit instruction on summarizing the main idea, predicting events and outcomes of upcoming text, drawing inferences, and other skills.

> Students should perform writing exercises every day to gain comfort and familiarity with writing. Instruction should be designed with the understanding that invented spelling does not conflict with teaching correct spelling, but can actually be helpful for developing understanding of the sounds that different combinations of letters create. Conventional spelling should be developed through focused instruction and practice, and primary-grade children should spell previously studied words correctly in their final writing products.

Children at Risk

Children who have successfully learned to read by elementary school have mastered three skills: They understand that letters of the alphabet represent word sounds, they are able to read for meaning, and they read fluently. Disruption of any of these components can throw off a child's development, the report says, and could lead to difficulties that ultimately will reduce the chances that the child will finish high school, get a job, or become an informed citizen.

Success in reading builds on the same complex set of skills for all children. Those running into difficulties do not need different instruction from other children, the report says, though they may need more focused, intense, and individual application of the same principles. Any special services they receive should be integrated into high-quality classroom instruction.

Reading problems are disproportionately high among minorities, non-English-speaking children, and those who grow up in poor or urban environments. A particularly thorny political problem has centered on how to educate children whose first language is not English. The report says that these children should first learn the skills of reading in their initial language -- the language in which they will best be able to discern the meaning of words and of sentences. If such instruction is not feasible in a given school system, the child should not be rushed prematurely into English reading instruction, but should be given an opportunity to develop a reasonable level of oral proficiency in English before learning to read. Children at risk of reading difficulties because of hearing impairment, language problems, or for other reasons must be identified quickly by pediatricians, social workers, and other early childhood practitioners.

To address these children's needs, the committee called for an increase in affordable, language-rich preschool programs. Programs designed as prevention for children at risk should focus on social, language, and cognitive development, not just on literacy. Organizations and government bodies concerned with the education of young children should target parents, care givers, and the general public in a campaign to promote public understanding of the way young children learn to read. The program should address ways of using books and creating

The program should address ways of using books and creating opportunities for building language skills and literacy growth through everyday activities.

Teacher Preparation

Because major responsibility for preventing reading difficulties is borne by early childhood educators and elementary school teachers, it is critical that they are sufficiently trained for the task. However, many teachers are not adequately prepared, the report says. Practitioners dealing with children under the age of eight need better training in reading development, and primary school teachers need ongoing professional development and continuing opportunities for mentoring and collaborating with reading specialists.

State certification requirements and teacher education curricula should be changed to incorporate key concepts about the way language relates to reading, as well as information about the relationship between early literacy behavior and conventional reading, the report says. Local school officials need to improve their staff development opportunities, which are often weakened by a lack of substantive, research-based content and systematic follow-up.

Schools that lack or have abandoned the use of reading specialists should re-examine their need for them and provide the functional equivalent of these well-trained staff members. These specialists' roles should be designed to ensure an effective two-way dialogue with regular classroom teachers. Volunteer tutors can be helpful in giving kids practice in reading for fluency, but are unlikely to be able to deal effectively with children who have serious reading problems.

A committee roster follows. The study was sponsored by the U.S. Department of Education and the Department of Health and Human Services. The National Research Council is the principal operating arm of the National Academy of Sciences and the National Academy of Engineering. It is a private, non-profit organization that provides advice on science and technology under a congressional charter granted to the National Academy of Sciences.

*Pre-publication copies of ***Preventing Reading Difficulties in Young Children*** are available from the National Academy Press at the mailing address in the letterhead; tel. (202) 334-3313 or 1-800-624-6242. The cost of the report is $45.00 (prepaid) plus shipping charges of $4.00 for the first copy and $.50 for each additional copy. Reporters may obtain a copy from the Office of News and Public Information at the letterhead address (contacts listed above).

NATIONAL RESEARCH COUNCIL
Commission on Behavioral and Social Sciences and Education
Division on Education, Labor, and Human Performance

1-22

CATCH THEM BEFORE THEY FALL

Identification and Assessment To Prevent Reading Failure in Young Children

BY JOSEPH K. TORGESEN

ONE OF the most compelling findings from recent reading research is that children who get off to a poor start in reading rarely catch up. As several studies have now documented, the poor first-grade reader almost invariably continues to be a poor reader (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Torgesen & Burgess, 1998). And the consequences of a slow start in reading become monumental as they accumulate exponentially over time. As Stanovich (1986) pointed out in his well-known paper on the "Matthew effects" (the rich get richer and the poor get poorer) associated with failure to acquire early word reading skills, these consequences range from negative attitudes toward reading (Oka & Paris, 1986), to reduced opportunities for vocabulary growth (Nagy, Herman, & Anderson, 1985), to missed opportunities for development of reading comprehension strategies (Brown, Palinscar, & Purcell, 1986), to less actual practice in reading than other children receive (Arlington, 1984).

The best solution to the problem of reading failure is to allocate resources for early identification and prevention. It is a tragedy of the first order that while we know clearly the costs of waiting too long, few school districts have in place a mechanism to identify and help children before failure takes hold. Indeed, in the majority of cases, there is no systematic identification until third grade, by which time successful remediation is more difficult and more costly.

School-based preventive efforts should be engineered to maintain growth in critical word reading skills at *roughly normal* levels throughout the early elementary school period. Although adequate development of these skills in first grade does not guarantee that children will continue to maintain normal growth in second grade without extra help, to the extent that we allow children to fall seriously behind at any point during early elementary school, we are moving to a "remedial" rather than a "preventive" model of intervention. Once children fall behind in the growth of critical word reading skills, it may require very intensive interventions to bring them back up to adequate levels of reading accuracy (Allington & McGill-Franzen, 1994; Vaughn & Schumm, 1996), and reading fluency may be even more difficult to restore because of the large amounts of reading practice that is lost by children each month and year that they remain poor readers (Rashotte, Torgesen, & Wagner, 1997).

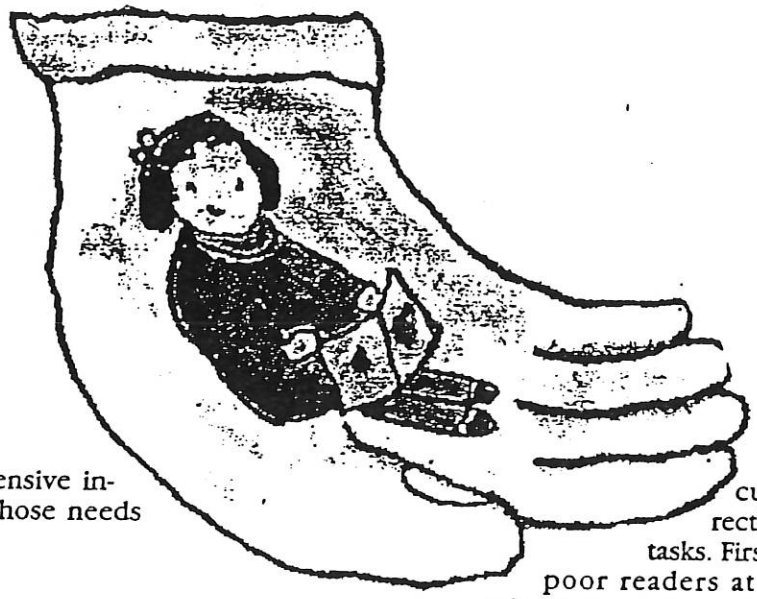
The purpose of this article is to provide practical advice about methods to prevent reading failure that is grounded in the new knowledge about reading we have acquired over the past two decades. My primary focus will be on early identification of children at risk for problems in learning to read as well as methods for monitoring the growth of critical early reading skills. The goal is to describe procedures that will allow educators to *identify children who need extra help in reading before they experience serious failure* and to monitor the early development of reading skill to *identify children who may require extra help as reading instruction proceeds* through elementary school.

The advice provided in this article is based on the research my colleagues Richard Wagner, Carol Rashotte, and I have been conducting on both prediction and prevention of reading disabilities (Torgesen, Wagner, & Rashotte, 1994; 1997; Wagner, et al., 1994; 1997) as well as the work of many other researchers that was reviewed in an earlier issue of this magazine (Summer, 1995). It is guided by several important assumptions and facts about reading, reading growth, and reading failure that will be discussed first. Following this description of assumptions and a brief outline of some critical dimensions of preventive instruction, I will describe a number of specific measures and procedures that should prove useful as educators seek

---

Joseph K. Torgesen is currently a Distinguished Research Professor of psychology and education at Florida State University. For the last ten years, he has been part of the research effort sponsored by the National Institutes of Health to identify the nature, causes, and best approaches to instruction for children with moderate to severe reading problems. *The research conducted at Florida State University that is cited in this article was supported by grants numbered HD23340 and HD30988 from the National Institute of Child Health and Human Development, and by grants from the National Center for Learning Disabilities and the Donald D. Hammill Foundation.*

1-23



ways to focus more intensive instruction on children whose needs are greatest.

Assumptions about reading, reading growth, and reading failure

Most of the points that will be discussed in this section are not, in fact, mere assumptions about reading, but, rather, are well-established facts. However, I use the word assumption here to convey the sense either that the ideas are self-evident or that they are now assumed to be true based on consistent research findings. The first of these "assumptions" is, in fact, a self-evident value judgment.

Adequate reading comprehension is the most important ultimate outcome of effective instruction in reading. The ultimate purpose of reading instruction is to help children acquire the skills that enable learning from, understanding, and enjoyment of written language. This "assumption" is not controversial. No matter what one's personal preferences for instructional method, the end goal is to help children comprehend written material at a level that is consistent with their general intellectual abilities.

Two general types of skill and knowledge are required for good reading comprehension. Consistent with Gough's "simple view of reading" (1996), comprehension of written material requires: 1) general language comprehension ability; and 2) ability to accurately and fluently identify the words in print. Knowledge and active application of specific reading strategies is also required to maximize reading comprehension (Mastropieri & Scruggs, 1997) but most of the variability among children and adults in comprehension of written material can be accounted for by measuring the two broad families of skills identified in Gough's simple view (Hoover & Gough, 1990). That is, good general language comprehension and good word reading skills are the most critical skills required for effective comprehension of written material.

Most children who become poor readers experience early and continuing difficulties in learning how to accurately identify printed words. This difficulty is expressed most directly on two kinds of reading tasks. First, children destined to be poor readers at the end of elementary school almost invariably have difficulties understanding and applying the alphabetic principle in deciphering unfamiliar words. These children have unusual difficulties learning to use the regular patterns of correspondence between letters and sounds in words as an aid in identifying new words they encounter in text (Siegel, 1989). They have trouble "sounding out" unknown words. Second, poor readers at all grade levels are characterized by slower than normal development of a "sight vocabulary" of words they can read fluently and automatically. Ultimately, it is this difficulty in rapid word recognition that limits comprehension in older poor readers, for these skills allow children to focus on constructing the meaning of what they are reading rather than spending too many of their intellectual resources on trying to identify the words (Adams, 1990). The strongest current theories of reading growth link phonetic and "sight word" reading skills together by showing how good phonetic reading skills are necessary in the formation of accurate memory for the spelling patterns that are the basis of sight word recognition (Ehri, in press; Share & Stanovich, 1995).

The most common cause of difficulties acquiring early word reading skills is weakness in the ability to process the phonological features of language (Liberman, Shankweiler, & Liberman, 1989). This is perhaps the most important discovery about reading difficulties in the last twenty years. Weaknesses in the phonological area of language development can be measured by a variety of nonreading tasks, but the ones most commonly used assess phonemic awareness, which can be defined simply as the ability to identify, think about, or manipulate the individual sounds in words. Much of our new confidence in being able to identify children at risk for reading failure before reading instruction begins depends on the use of tests of phonemic awareness, since this ability has been shown to be causally related to the growth of early word reading skills (Lundberg, Frost, & Peterson, 1988; Wagner, et al., 1997).

Discovery of the core phonological problems associ-

1-24

with specific reading disability has had at least one unanticipated consequence. The ability to assess these core language problems directly has led to the discovery that the early word reading difficulties of children with relatively low general intelligence and verbal ability are associated with the same factors (weaknesses in phonological processing) that interfere with early reading growth in children who have general intelligence in the normal range (Fletcher, et al., 1994; Share & Stanovich, 1995; Stanovich & Siegel, 1994). So, weaknesses in phonemic awareness characterize children with reading problems across a broad span of general verbal ability. On the one hand, many children enter school with *adequate general verbal ability* and cognitive weaknesses limited to the phonological/language domain. Their primary problem in learning to read involves learning to translate between printed and oral language. On the other hand, another significant group of poor readers, composed largely of children from families of lower socio-economic or minority status, enter school significantly delayed in a much broader range of prereading skills (Whitehurst & Lonigan, in press). Since these children are delayed not only in phonological but also in general oral language skills, they are deficient in both of the critical kinds of knowledge and skill required for good reading comprehension. Even if these children can acquire adequate word reading skill, their ability to comprehend the meaning of what they read may be limited by their weak general verbal abilities.

Children with general oral language weaknesses require extra instruction in a broader range of knowledge and skills than those who come to school impaired only in phonological ability. What is well established at this point, though, is that *both kinds of children* will require special support in the growth of early word reading skills if they are to make adequate progress in learning to read.

Elements of an effective preventive program in reading

The most critical elements of an effective program for the prevention of reading disability at the elementary school level are: (a) the right kind and quality of instruction delivered with the (b) right level of intensity and duration to (c) the right children at the (d) right time. I will briefly consider each of these elements in turn.

The right kind and quality of instruction. It is beyond the scope of this article to discuss instructional methods for children with phonological processing weaknesses in any depth at all. In broad stroke, they will benefit from the same approach to reading instruction as children with normal abilities in this area—structured, systematic, and explicit—but for this at-risk group, such instruction is not just beneficial, it is critical. As experienced teachers understand (Gaskins, et al., 1996), we cannot assume that these children will acquire any necessary skill for reading words unless they are directly taught that skill or knowledge and receive sufficient opportunities to practice it. Some of the word-level skills and knowledge these children will require instruction on include: phonemic awareness, letter-sound correspondences, blending skills, small number of pronunciation conventions (i.e., silent *e* rule), use of context to help specify a word once it is partially or completely phonemically decoded, strategies for multi-syllable words, and automatic recognition of high-frequency "irregular" words. It goes almost without saying that this type of instruction should be embedded within as many opportunities for meaningful reading and writing as possible.

The lesson from recent large-scale prevention studies (Brown & Felton, 1990; Foorman, et al., 1998; Torgesen, et al., 1998; Vellutino, et al., 1997) is that it is possible to maintain critical word reading skills of most children at risk for reading failure at roughly average levels if this type of instruction is provided beginning sometime during kindergarten or first grade. However, it is also true that in all studies conducted to date, substantial proportions of children with the most severe weaknesses remain significantly impaired in these critical skills following intervention. For example, if we adopt the 30th percentile as a standard for adequate reading progress, then the proportion of the total population remaining at risk in spite of the best interventions tested to date ranges from 5 percent to 7 percent (Torgesen, 1998). Although these results are clearly better than the 30 percent to 60 percent of children who frequently fall below these standards without special interventions, they nevertheless suggest that there is a core of disabled readers in the population for whom we have not yet solved the reading puzzle.

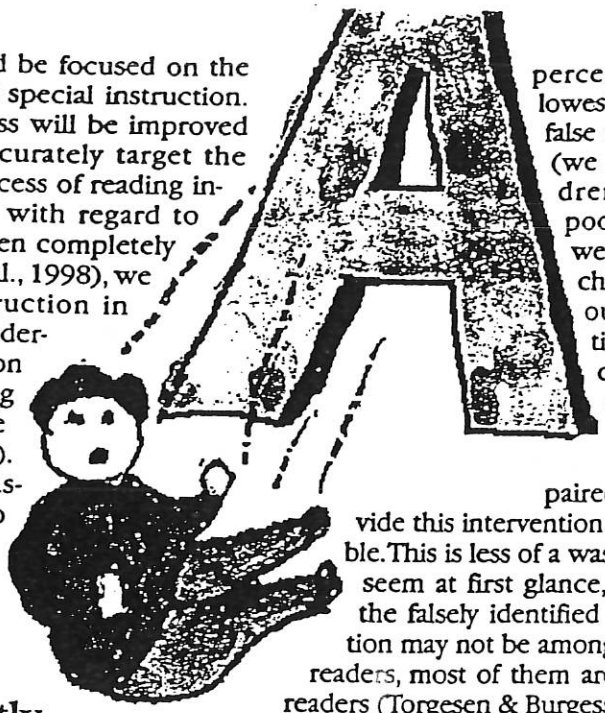
It is almost certain that some additional answers to this question will come as we direct our attention to the quality and intensity, as well as the content, of our instruction. For example, Juel (1996) has shown the importance of a particular kind of "scaffolded" interaction between teacher and child in increasing understanding and application of phonemic reading skills, and these types of interactions are also prescribed in the teacher manuals of at least two widely used instructional programs designed for children with reading disabilities (Lindamood & Lindamood, 1984; Wilson, 1988). We turn now to a brief consideration of issues surrounding intensity of instruction.

The right level of intensity. Greater intensity and duration of instruction is required because the increased explicitness of instruction for children who are at risk for reading failure requires that more things be taught directly by the teacher. Intensity of instruction is increased primarily by reducing teacher/student ratios. Unless beginning reading instruction for children with phonological weaknesses is more intensive (or lasts significantly longer) than normal instruction, these children will necessarily lag significantly behind their peers in reading growth. An effective preventive program may involve several levels of instructional intensity ranging from small-group to one-on-one instruction, depending upon the severity of the risk factors for each child.

The right children at the right time. These factors are considered together because they are both tied directly to the availability of accurate identification procedures at various age levels. That is, to be most effi-

1-25

ient, a preventive program should be focused on the children who are most in need of special instruction. The efficiency of the entire process will be improved if procedures are available to accurately target the right children very early in the process of reading instruction. Although timing issues with regard to preventive instruction have not been completely resolved by research (Torgesen, et al., 1998), we do know, for example, that instruction in phonological awareness during kindergarten can have a positive effect on reading growth after formal reading instruction begins in the first grade (Lundberg, Frost, & Peterson, 1988). Thus, I have proceeded on the assumption that it will be useful to identify high-risk children at some time during the kindergarten year so that preventive work may begin as early as possible.



How accurate are currently available early identification procedures?

As stated earlier, the primary purpose of this article is to make some practical suggestions about procedures and tests that can be used to identify children for preventive reading or prereading instruction. From the outset, however, it is important to recognize that our ability to predict which children will have the most serious reading difficulties is still far from perfect. For example, in a recent comprehensive review of early identification research (1998), Scarborough pointed out that all studies continue to report substantial levels of two kinds of prediction errors.

False positive errors are made when children who will eventually become good readers score below the cut-off score on the predictive instrument and are falsely identified as "at risk." In general, the proportion of this type of error has ranged between 20 percent and 60 percent, with an average of around 45 percent. That is, almost half of the children identified during kindergarten as "at risk" turn out not to have serious reading problems by the end of first grade. *False negative* errors occur when children who later exhibit reading problems are identified as not being at risk. Typical percentages of false negative errors range from 10 percent to 50 percent, with an average of around 22 percent. That is, on average, current procedures fail to identify about 22 percent of children who eventually end up with serious reading difficulties.

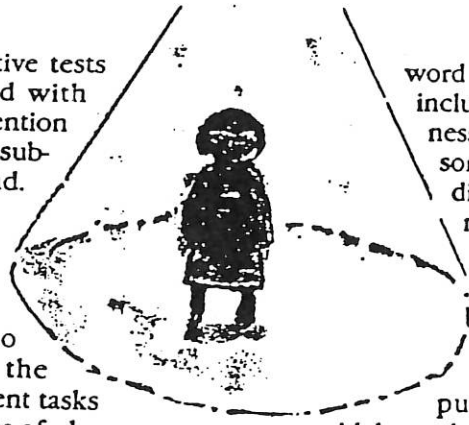
In any given study, the relative proportion of false positive and false negative errors is somewhat arbitrary, since it depends on the level of the cut-off score. For example, we reported a significant reduction in the percentage of false negative errors within the same sample of children by doubling the number of children we identified as at risk (Torgesen, in press; Torgesen & Burgess, 1998). Our goal was to identify, during the first semester of kindergarten, the children most at risk to be in the bottom 10 percent in word reading ability by the beginning of second grade. When we selected the 10 percent of children who scored lowest on our predictive tests, our false negative rate was 42 percent (we missed almost half the children who became extreme poor readers). However, when we identified the 20 percent of children who scored lowest on our measures, the false negative rate was reduced to 8 percent. As a practical matter, if schools desire to maximize their chances for early intervention with the most impaired children, they should provide this intervention to as many children as possible. This is less of a waste of resources than it might seem at first glance, because, although many of the falsely identified children receiving intervention may not be among the most seriously disabled readers, most of them are likely to be below-average readers (Torgesen & Burgess, 1998).

Two other pieces of information are relevant to the selection of procedures for early identification of children at risk for reading difficulties. First, prediction accuracy increases significantly the longer a child has been in school. Prediction of reading disabilities from tests given at the beginning of first grade is significantly more accurate than from tests administered during the first semester of kindergarten (Scarborough, 1998; Torgesen, Burgess, & Rashotte, 1996). Given the widely varying range of children's preschool learning opportunities, many children may score low on early identification instruments in the first semester of kindergarten simply because they have not had the opportunity to learn the skills. However, if prereading skills are actively taught in kindergarten, some of these differences may be reduced by the beginning of the second semester of school. Thus, I would recommend that the screening procedures described here not be administered until the beginning of the second semester of kindergarten, at which time they will be much more efficient in identifying children who will require more intensive preventive instruction in phonemic awareness and other early reading skills.

Second, although batteries containing multiple tests generally provide better prediction than single instruments, the increase in efficiency of multi-test batteries is generally not large enough to warrant the extra time and resources required to administer them (Scarborough, 1998). Thus, I recommend an identification procedure involving administration of two tests: 1) a test of knowledge of letter names or sounds; and 2) a measure of phonemic awareness. Measures of letter knowledge continue to be the best single predictor of reading difficulties, and measures of phonemic awareness contribute additional predictive accuracy. In our experience, tests of letter name knowledge are most predictive for kindergarten children, and tests of letter-sound knowledge are most predictive for first graders. Since reading growth is influenced by noncognitive factors such as attention/motivation and home background (Torgesen, et al., 1998), as well as specific knowledge

1-26

skills, scores from these objective tests might profitably be supplemented with teacher ratings of behavior and attention to identify children most at risk for subsequent difficulties in learning to read.



How should phonemic awareness be assessed?

Since researchers first began to study phonological awareness in the early 1970s, more than twenty different tasks have been used to measure awareness of phonemes in words. These measures can be grouped into three broad categories: sound comparison, phoneme segmentation, and phoneme blending.

■ *Sound comparison* tasks use a number of different formats that all require children to make comparisons between the sounds in different words. For example, a child might be asked to indicate which word (of several) begins or ends with the same sound as a target word (i.e., "Which word begins with the same first sound as *cat: boy, cake,* or *fan*?"). Additionally, tasks that require children to generate words that have the same first or last sound as a target word would fall in this category. Sound comparison tasks are among the least difficult measures of phonemic awareness, and thus are particularly appropriate for kindergarten age children.

■ *Phoneme segmentation* tasks involve counting, pronouncing, deleting, adding, or reversing the individual phonemes in words. Common examples of this type of task require pronouncing the individual phonemes in words ("Say the sounds in *cat* one at a time."), deleting sounds from words ("Say *card* without saying the /d/ sound."), or counting sounds ("Put one marker on the line for each sound you hear in the word *fast.*")

■ *Phoneme blending* skill has only been measured by one kind of task. This is the sound-blending task in which the tester pronounces a series of phonemes in isolation and asks the child to blend them together to form a word (i.e., "What word do these sounds make, /f/ - /a/ - /t/?"). Easier variants of the sound-blending task can be produced by allowing the child to choose from two or three pictures the word that is represented by a series of phonemes.

In general, these different kinds of phonemic awareness tasks all appear to be measuring essentially the same construct, or ability. Although some research (Yopp, 1988) has indicated that the tasks may involve different levels of intellectual complexity, and there may be some differences between segmentation and blending tasks at certain ages (Wagner, Torgesen, & Rashotte, 1994), for the most part, they all seem to be measuring growth in the same general ability (Hoien, et al., 1995; Stanovich, Cunningham, & Cramer, 1984). Sound comparison measures are easiest and are sensitive to emergent levels of phonological awareness, while segmentation and blending measures are sensitive to differences among children during later stages of development involving refinements in explicit levels of awareness. Measures of sensitivity to rhyme ("Which word rhymes with *cat: leg* or *mat*[illegible] not included as measures of phonemic awareness because they appear to be measuring something a little different, and less predictive of reading disabilities, from those measure that ask children to attend to individual phonemes. For the same reason, measures of syllable awareness are not included in this group.

Measures of phonemic awareness that are suited for early identification purposes include the following three widely used tests:

The Phonological Awareness Test (Robertson & Salter, 1995). This test contains five different measures of phonemic awareness, plus a measure of sensitivity to rhyme. The five measures of phonemic awareness are segmentation of phonemes, phoneme isolation, phoneme deletion, phoneme substitution, and phoneme blending. The phoneme isolation test, which requires children to pronounce the first, last, or middle sounds in words, would appear to have the most appropriate level of difficulty for kindergarten screening (the test should be easy enough so that only the most delayed children will do poorly on it), and any of the others could be used for first- or second-grade assessments. *The Phonological Awareness Test* is nationally normed on children from age five through nine, and it can be ordered from LinguiSystems, 3100 4th Avenue, East Moline, IL 61244-0747. Phone: 800-776-4332. The cost of a test manual, test supplies, and fifteen test booklets is $69.

The Test of Phonological Awareness (Torgesen & Bryant, 1994). This test was designed as a group-administered test of phonemic awareness for kindergarten and first-grade children. It was specifically constructed to be most sensitive to children with weaknesses in development in this area, which helps make it appropriate for identifying at-risk children. The kindergarten version of the test requires children to notice which words (represented by pictures) begin with the same first sound, while the first-grade version asks them to compare words on the basis of their last sounds. It can be easily administered to groups of five to ten children at a time. *The Test of Phonological Awareness* is nationally normed, and it can be ordered from PRO-ED Publishing Company, 8700 Shoal Creek Blvd., Austin, TX 78757-6897. Phone: (512) 451-3246. The cost of a test manual and a supply of fifty test forms (twenty-five kindergarten version, twenty-five elementary school version) is $124.

The *Yopp-Singer Test of Phoneme Segmentation* (Yopp, 1995) is a brief test of children's ability to isolate and pronounce the individual phonemes in words. This is a task that has been widely used in research on phoneme awareness over the past twenty years, and it is highly correlated with other measures of phoneme awareness. The test was designed for children in kindergarten, but it should also be appropriate for identifying children who are weak in phonemic awareness during first grade. The test has twenty-two items that are all of the same type and that ask the child to pronounce each of the phonemes in words that vary from two to three phonemes in length. The test does

1-27

have norms with it, but it is available free in volume 49 (1995) of the widely read journal *The Reading Teacher*, pp. 20-29.

The measurement of letter knowledge

In all of our research, we have measured letter knowledge in two ways. We measure *letter name* knowledge by presenting each letter in simple upper-case type on a single card and asking for its name. The score on this test is simply the number of letters for which the child can give the appropriate name. We measure *letter-sound* knowledge by presenting all letters in lower-case type and asking for the "sound the letter makes in words." If a consonant letter can commonly represent two different sounds (i.e., c, g) we probe for the second sound, and we also ask for the long and short pronunciation of each vowel. The score is the total number of sounds the child can give. We have found that letter-name knowledge is a more sensitive predictor for kindergarten children, while letter-sound knowledge is a better predictor for children in first grade. Two tests that provide nationally standardized norms for performance on letter-name and letter-sound knowledge are:

The *letter identification* subtest of the *Woodcock Reading Mastery Test-Revised* (Woodcock, 1987). This test does not measure simple letter-name knowledge in the way we assess it, because it presents letters in several different fonts, some of which may be unfamiliar to children. It also allows children to give either the name or the sound the letter makes in words. However, children who perform poorly in kindergarten (do not know the names of very many letters) will not reach the more difficult items, so that their score should be quite comparable to a more straightforward test of letter-name knowledge. *The Reading Mastery Test-Revised* is available from American Guidance Service, 4201 Woodland Road, Circle Pines, MN 55014-1796. Phone (800) 328-2560. The cost for the manual and forms is $314.95.

The *graphemes* subtest of the *Phonological Awareness Test* (Robertson & Salter, 1995). This test provides a comprehensive assessment of letter-sound knowledge extending from single consonants (i.e., b,c,k,m) through vowel digraphs and diphthongs (i.e., ea, ai, ow, oy). As mentioned before, it is standardized on children from aged five through nine.

Is it necessary for a test to be nationally standardized for it to be useful in early identification?

This issue is important because of the potential expense of employing standardized measures in large-scale screening efforts. Nationally based norms are *not* required to identify which children within a given classroom or school are weakest in phonemic awareness and letter knowledge. However, the proportion of children who come to school with weak skills and knowledge in these areas will depend somewhat on specific aspects of their preschool language and literacy environment and will almost certainly vary from school to school across different communities. Tests with national norms can help to pinpoint classes or schools in which a special effort must be made to enhance phonological awareness in children prior to, or during, reading instruction. For example, a classroom in which 75 percent of the children performed below the 20th percentile (in the bottom 20 percent of all children), will require more instructional resources to prepare children for learning to read than a classroom in which only 10 percent of the children scored that low. Without norms, it is possible to identify weak children within a given class or school, but it is not possible to determine what proportion of children in the entire school may require intervention because of relatively weak prereading skills and knowledge. On the one hand, if classroom resources allow extra help for only a fixed number of children (say, 20 percent to 30 percent), then measures without national norms can be used to identify the group of children within the classroom most in need of intervention. On the other hand, if the goal is to determine the amount of resources that may be needed to help all children with relatively weak skills in these areas, then normative measures will be required.

The combination of letter knowledge and phonemic awareness tests I have recommended should take no more than ten to fifteen minutes per child to administer. The tests do not require highly trained personnel to administer them, although anyone who tests young children must be very familiar with the tests and be able to establish a supportive rapport.

Monitoring growth in early reading skills

Once reading instruction begins, the best predictor of future reading growth is current reading achievement, and the most critical indicators of good progress in learning to read during the early elementary period are measures of word reading skill. Children who end up as poor readers at the end of elementary school are almost invariably those who fail to make normal progress in these skills during the first years of elementary school. These children are most frequently impaired in both the ability to apply phonetic strategies in reading new words and in the ability to retrieve sight words from memory. They not only have difficulty becoming accurate in the application of these processes but also they frequently have special difficulties with becoming fluent in their application. Before discussing specific methods for the diagnostic assessment of these word reading skills, one general issue regarding reading assessment requires discussion.

First, the assessment that will be recommended here is very different from the "authentic literacy assessment" that is currently advocated by many reading professionals (Paris, et al., 1992). Authentic assessment is different in at least two ways from the reading assessment measures we will be discussing. First, the goal of "authentic assessment" is to measure children's application of broad literacy skills to authentic tasks, like gathering information for a report, use of literacy as a medium for social interactions, or ability to read a selec-

1-28

nd then write a response to it. It also seeks to re children's enjoyment, ownership, and involvement in literacy activities both at school and at home.

This kind of assessment is a clear complement to the type of assessments we will describe for monitoring growth in word level reading skills. All of the literacy outcomes that are part of authentic assessment are important parts of a total literacy assessment program. After all, if a child can read, but does not enjoy reading and does not apply important literacy skills to everyday tasks, then some important goals of literacy instruction have not been attained.

However, since these procedures are focused on high-level reading outcomes, they cannot provide precise information about level of performance on important subskills in reading. If a child's overall performance on authentic literacy tasks is limited, it is frequently difficult to obtain a precise estimate of the specific component skills that are weak. The goal of the kind of assessments we will discuss here is to quantify the degree of skill a child possesses in word identification processes that have been shown to be a critical foundation for overall reading success.

Commonly used diagnostic measures of word reading ability

It is beyond the scope of this article to identify all the available tests of word level reading skills. Rather, I will provide examples of measurement strategies from the most commonly used measures.

Sight word reading ability. Two measures are widely used in this area, and both involve the same assessment strategy. The Word Identification subtest from the *Woodcock Reading Mastery Test-Revised* (Woodcock, 1987), and the reading subtest of the *Wide Range Achievement Test-3* (Wilkinson, 1995) both require children to read lists of words that gradually increase in length and complexity while decreasing in frequency of occurrence in printed English. For example, the easiest three words on the Word Identification subtest are *go, the,* and *me,* words of mid-level difficulty are *pioneer, inquire,* and *wealth,* and the hardest three are *epigraphist, facetious,* and *shillelagh.*

Neither of these widely used tests place stringent time pressure on students, so both phonetic decoding processes and sight word processes can be used to identify words on these lists. Both tests have been normed nationally, and one of their strengths is that they allow a direct assessment of children's ability to identify words solely on the basis of the word's spelling. When reading text, children also have context clues available to assist word identification, and thus text-based measures, although they may be more "authentic" in one sense, are less direct in their assessment of the kinds of word-processing skills that are particularly deficient in children with reading problems.

Phonetic reading ability. The single best measure of children's ability to apply knowledge of letter-sound correspondences in decoding words is provided by measures of nonword reading (Share & Stanovich, 1995). The Word Attack subtest of the *Woodcock Reading Mastery Test-Revised* (Woodcock, 1987) is a good example of this kind of diagnostic test. It co a series of increasingly complex nonwords that en are asked to "sound out as best they can." The three easiest items on the test are *ree, ip,* and *din;* items of moderate difficulty are *rejune, depine,* and *viv;* and the three hardest items are *pnir, ceisminadolt,* and *byrcal.* Because the words are presented out of context, they stress the child's ability to fully analyze each word to produce the correct pronunciation. On the other hand, measures such as this do not allow an assessment of children's ability to combine phonetic decoding with use of context to arrive at a word's correct pronunciation. However, since both good and poor readers appear able to use context equally well (as long as the context is understood, Share & Stanovich, 1995), this is not an important omission on a diagnostic measure of word reading ability.

Word reading fluency. Word reading fluency measures have typically measured rate of reading connected text. One of the more widely used measures in this area is the *Gray Oral Reading Test-3rd Edition.* (Wiederholt & Bryant, 1992). This test consists of thirteen increasingly difficult passages, each followed by five comprehension questions. A measure of oral reading rate is obtained by recording the time it takes for the child to read each passage. One potential problem with the Gray Oral Reading Test is that it does not provide a very sensitive measure of individual differences in word reading ability at very low levels of performance, such as those found in beginning first graders, or disabled readers through second grade. The passages simply begin at too high a level for children with very poor or undeveloped reading skills to display the word reading skills they actually possess.

In an effort to provide measures of fluency and accuracy in word reading skill that are simple to administer and sensitive to individual differences across a broad range of reading skills, we are currently developing simple measures of *Word Reading Efficiency* and *NonWord Efficiency* (Torgesen & Wagner, 1997). In both of these measures, children are shown lists of increasingly difficult words and nonwords and asked to read as many words as possible in forty-five seconds. There are two forms to each test, and the child's score is simply the average number of words read in forty-five seconds. Initial evaluations indicate that these measures are very reliable (parallel form reliabilities vary between .97 and .98 for kindergarten through fifth grade). They are also highly correlated with corresponding measures from the *Woodcock Reading Master Test-Revised* at early grades (when children often run out of words they can read before they run out of time, correlations range from .89 to .94) and slightly less correlated (.86 to .88) at fourth grade, when fluency of word reading processes becomes more important to performance on the tests. These tests have been standardized nationally and will be available from PRO-ED publishing company in late summer 1998. If a single form of each test is administered, it will provide indices of growth in phonetic decoding and sight word reading that can be administered several times during the year and that take a very short amount of time to give.

1-29

To summarize, adequate monitoring of the growth of children's word reading abilities should include out-of-context measures of word reading ability, phonetic decoding ability (as measured by ability to read non-words), and word reading fluency. The fluency measures become more important after about second to third grade, when children have acquired a fund of word reading skills they can apply with reasonable accuracy. Measures that involve out-of-context word reading more directly assess the kinds of word reading skills that are particularly problematic for children with reading disabilities because they eliminate the contextual support on which these children rely heavily. To obtain a *complete* picture of overall reading development, however, it is also important to observe the way that the child integrates all sources of information about words in text, and this can only be estimated by carefully observing children as they read connected passages. □

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Allington, R.L. (1984). Content coverage and contextual reading in reading groups. *Journal of Reading Behavior,* 16, 85-96.

Allington, R.L., & McGill-Franzen, A. (1994). Reading and the mildly handicapped. *International Encyclopedia of Education,* Oxford, U.K.: Pergamon.

Brown, A. L., Palincsar, A. S., & Purcell, L. (1986). Poor readers: Teach, don't label. In U. Neisser (Ed.) *The School Achievement of Minority Children: New Perspectives,* (pp. 105-143) NJ: Lawrence Erlbaum Associates.

Ehri, L.C. (in press). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala & L. Ehri (Eds.). *Word recognition in beginning reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I.Y., Stuebing, K.K., Francis, D. J., Fowler, A. E., & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology,* 86, 6-23.

Foorman, B.R., Francis, D.J., Fletcher, J.M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology.*

Francis, D.J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology,* 88, 3-17.

Gaskins, I.W., Ehri, L.C., Cress, C., O'Hara, C., & Donnelly, K. (1996). Procedures for word learning: Making discoveries about words. *The Reading Teacher,* 50, 312-327.

Gough, P. B. (1996). How children learn to read and why they fail. *Annals of Dyslexia,* 46, 3-20.

Hoien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal,* 7, 171-188.

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing,* 2, 127-160.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly,* 31, 268-289.

Liberman, I.Y., Shankweiler, D., & Liberman, A.M. (1989). The alphabetic principle and learning to read. In Shankweiler, D. & Liberman, I.Y. (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor, MI: U. of Michigan Press.

Lindamood, C. H., & Lindamood, P. C. (1984). *Auditory Discrimination in Depth.* Austin, TX: PRO-ED, Inc.

undberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in pre-school children. *Reading Research Quarterly;* 23, 263-284.

Mastropieri, M.A. & Scruggs, T.E. (1997). Best practices in promoting reading comprehension in students with learning disabilities: 1976-1996. *Remedial and Special Education,* 18, 197-213.

Nagy, W.E., Herman, P.A., & Anderson, R.C. (1985). Learning words from context. *Reading Research Quarterly,* 20, 233-253.

Oka, E., & Paris, S. (1986). Patterns of motivation and reading skills in underachieving children. In S. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2). Hillsdale, NJ: Erlbaum.

Rashotte, C.A., Torgesen, J.K., & Wagner, R.K. (1997). *Growth in reading accuracy and fluency as a result of intensive intervention.* Paper presented at the annual meetings of the Florida Branch of the International Dyslexia Association, Miami, FL, September.

Scarborough, H.S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In Shapiro, B.K., Accardo, P.J., & Capute, A.J. (Eds.) *Specific Reading Disability: A view of the spectrum.* (pp. 75-120). Timonium, MD: York Press, Inc.

Share, D.L., & Stanovich, K.E. (1995). Cognitive processes in early reading development: A model of acquisition and individual differences. *Issues in Education: Contributions from Educational Psychology,* 1, 1-57.

Siegel, L.S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities,* 22, 469-479.

Stanovich, K.E., & Siegel, L.S. (1994). The phenotypic performance profile of reading-disabled children: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology,* 86, 24-53.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly,* 21, 360-406.

Stanovich, K.E., Cunningham, A.E., & Cramer, B.B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology,* 38, 175-190.

Torgesen, J.K. (in press). Assessment and instruction for phonemic awareness and word reading skills. In H.W. Catts and A.G. Kamhi (Eds.) *Language Basis of Reading Disabilities.* Needham Heights, MA: Allyn & Bacon.

Torgesen, J.K. (1998) *Individual differences in response to reading intervention.* Paper presented at the Pacific Coast Research Conference, LaJolla, CA, February.

Torgesen, J.K., & Burgess, S.R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. Metsala & L. Ehri (Eds.). *Word Recognition in Beginning Reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Torgesen, J.K., Burgess, S., & Rashotte, C.A. (1996). *Predicting phonologically based reading disabilities: What is gained by waiting a year?* Paper presented at the annual meetings of the Society for the Scientific Study of Reading, New York, April.

Torgesen, J.K.; Wagner, R.K., & Rashotte, C.A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities,* 27, 276-286.

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (1997). The prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading,* 1, 217-234.

Vaughn, S. & Schumm, J.S. (1996) Classroom ecologies: Classroom interactions and implications for inclusion of students with learning disabilities (pp. 107-124). In D. L. Speece & B.K. Keogh (Eds.), *Research on classroom ecologies.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1994). The development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology,* 30, 73-87.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., Hecht, S.A., Barker, T.A., Burgess, S.R., Donahue, J., & Garon, T. (1997). Changing causal relations between phonological processing abilities and word-level reading as children develop from beginning to fluent readers: A five-year longitudinal study. *Developmental Psychology,* 33, 468-479.

Whitehurst, G.J. & Lonigan, C.J. (in press). Child development and emergent literacy. *Child Development,* 69.

Wilson, B.A. (1988). *Instructor manual.* Millbury, MA: Wilson Language Training.

Yopp, H.K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly,* 23, 159-177.

1-30

EARLY CHILDHOOD EDUCATION

4-YEAR-OLD PRESCHOOL AND ALL-DAY KINDERGARTEN

Research Summary

Getting children off to a strong start means they will finish better off in the future. Studies show that is particularly true for educating our youth. Children who master basic skills by the end of 3rd grade/beginning of 4th grade are most likely to finish high school, avoid teen pregnancy, stay out of crime and off of welfare – they also have higher earnings and increase the tax base.

This sort of success starts in preschool and all-day Kindergarten. Studies show one year of preschool before starting Kindergarten pays off in the long run for all children, but especially for disadvantaged students. As adults, the at-risk students exposed to preschool make more money, have a higher percentage of home ownership and receive fewer social services than their counterparts without early learning. Those results pay off for the individual and for society as state spending on prisons, welfare and social services drop dramatically while the tax base increases. In fact, one study found preschool programs return an estimated $7.16 for every dollar spent. All-day Kindergarten has equally impressive results. Children excel in oral language development, math concepts and reasoning, beginning reading skills, and student behavior with benefits lasting well into the second grade. Preschool, coupled with all-day Kindergarten, stops social problems before they start – and gives children their best shot at a bright future.

Glasscock-Tanner "Years of Promise -- Kansas' Keys to Success" Education Plan

This plan focuses on bringing these individual and societal successes to a generation of young Kansas children, and every generation after them. Since preschool generates dramatic results with students in poverty, the plan calls for:

- Kansas elementary schools must offer a half-day early learning program for at-risk 4-year-olds to be utilized by parent choice.

- Kansas Kindergarten classes will be extended to full days for all children whose parents would like the opportunity for their child. The state will fund the extra half-day of Kindergarten for at-risk students. School districts may determine how to fund the additional classroom time for students who are not at-risk for failure. The districts will have the option to charge attendance fees for the cost of the additional half-day from parents of students who are not at-risk

By targeting our resources on the early school years, Kansas schools can give children the keys to success and prepare our youth for the challenges in life. A strong start for all students means a better future for all Kansans.

EARLY CHILDHOOD EDUCATION RESEARCH HIGHLIGHTS

"Children fortunate enough to attend a high-quality preschool ... and who enter the primary grades with adequate preparation have a better chance of achieving to high levels than those who do not."

Years of Promise: A Comprehensive Learning Strategy for America's Children
Carnegie Corporation of New York

"The estimated rate of return on preschool education exceeds the average rate of return in the stock market over the past 30 years."

"Programs that produce substantial improvements in...school success of children in poverty can be expected to produce substantial direct benefits through educational cost-savings and substantial indirect benefits as the result of increased productivity and social responsibility." *W. Steven Barnett*
"Long-Term Cognitive and Academic Effects of Early Childhood Education on Children in Poverty"
Preventive Medicine

"Compared with the preschool program's cost... [the] benefits make the program a worthwhile investment for taxpayers as well as for society in general. Over the lifetimes of the participants, the preschool program returns to the public an estimated $7.16 for every dollar invested."

"At age 27...the [preschool] program group had significantly higher monthly earnings...higher percentages of home ownership and second-car ownership...higher level of school completed... lower percentage receiving social services... fewer arrests .... and significantly fewer of the births to program females were out of wedlock."

L.J. Schweinhart, H.V. Barnes, and D.P. Weikart
"Executive Summary: The High/Scope Perry Preschool Study Through Age 27"
Significant Benefits: The High/Scope Perry Preschool Study Through Age 27

"When deciding what kind of kindergarten program to offer, educators must consider not only the desires of parents but also the potential effects on student achievement.... The results of statistical analyses conducted for ... full-day versus half-day comparison in oral language development, early mathematics concepts, and emergent literacy skills all indicated significant differences in favor of full-day kindergarten."

Y. Lawrence Wang and Whitcomb G. Johnstone
"Evaluation of a Full-Day Kindergarten Program"
ERS Spectrum

"Children who attend preschool prior to kindergarten experience greater subsequent success in elementary school than those who do not .... [and] participation in full-day kindergarten is positively related to subsequent school performance.... The benefits seem to last well into the second grade."

John Cryan, Robert Sheehan, Jane Wiechel and Irene Bandy-Hedden
"Success Outcomes of Full-Day Kindergarten: More Positive Behavior and Increased Achievement in the Years After"
Early Childhood Research

VE MEDICINE 27, 204–207 (1998)
ARTICLE NO. PM970275

Long-Term Cognitive and Academic Effects of Early Childhood Education on Children in Poverty

W. Steven Barnett, Ph.D.[1]

Rutgers–The State University of New Jersey, New Brunswick, New Jersey 08903

It is generally accepted that early childhood education improves the cognitive performance of children in poverty in the short-term, but whether cognitive effects persist in the long-term is hotly debated. This paper presents the results of a critical review of 38 studies of the long-term effects of early childhood programs on children in poverty. Outcomes examined include IQ, achievement, and academic success as measured by grade repetition, special education placement, and high school graduation. Early childhood education is found to produce persistent effects on achievement and academic success, but not on IQ (with some exceptions). Head Start and public school programs produce the same types of effects as better funded model programs, but at least some of the effects are smaller. Cost–benefit analysis based on one randomized trial finds that the economic return from providing early education to children in poverty far exceeds the costs. Head Start, public school preschool education, and education in high-quality child care programs all offer avenues for government investment to improve the long-term cognitive development and academic success of children in poverty.

Key Words: **cognitive development; achievement; academic success; special education; early childhood; preschool education; child care; cost–benefit analysis.**

Three questions are addressed in this paper. What are the long-term effects of early childhood education on the cognitive development and academic success of children in poverty? What are the economic consequences of these effects? What new public policies ought to be implemented based on our knowledge of long-term effects and their economic consequences?

[1] To whom correspondence and reprint requests should be addressed at Rutgers Graduate School of Education, 10 Seminary Place, New Brunswick, NJ 08903. Fax: (908) 932-6803.

LONG-TERM EFFECTS

The short-term effects of early childhood education on the cognitive development of children in poverty are well established. A wide range of programs specifically designed to improve the education of these children before age 5 have been shown to produce immediate effects on IQ and achievement of about 0.5 standard deviations, equivalent to about 8 IQ points [*1,2*]. Similar, though somewhat smaller, effects have been found for ordinary child care programs, which on average have less qualified and well-trained staff, larger class sizes, and less parental involvement and do not provide the same quality of educational experiences [*3,4*].

There has been less agreement about long-term effects, but the most common conclusion has been that effects on cognitive development decline after children leave the programs and are eventually lost altogether, while some effects may persist on measures of school success such as grade repetition and special education placements [*5*]. In addition, some have concluded that Head Start, public school, and other large-scale government efforts may not be able to reproduce the results of high-priced model programs operated by universities. Yet, the results of some studies are at odds with these conclusions, and the question arises as to why long-term effects on school success should persist if cognitive effects do not.

A review of the literature was conducted to address these issues. Thirty-eight studies that estimated effects of early childhood education programs (before age 5) on the cognitive development or school success of children in poverty at least through grade 3 were examined [*3,6*]. These included 15 studies of research-sponsored model early childhood programs. Most model programs were center-based, though some used home-visiting alone or together with a center-based approach. A few provided full-day child care. The other 23 studies investigated the effects of large-scale public programs provided by Head Start or the public schools. These were primarily half-day preschool education programs, though Head Start also provides medical and dental exams and care,

0091-7435/98 $25.00

immunizations, and parent activities including education and counseling.

IQ Effects

The general pattern was for effects on IQ test scores to fade out after entry to elementary school, in most cases quite rapidly. To some extent this appears to occur because children in poverty who do not receive early education outside the home experience an increase in IQ as a result of public education beginning in kindergarten. Two experimental studies that provided full-day educational child care programs from the first year of life through age 5 differ from other studies in that they find some effects on IQ (0.3 standard deviations, 5 points) persisting well into adolescence. This suggests that intensive programs from birth through age 5 might produce more persistent effects on IQ than part-day programs beginning at age 3 or 4.

Achievement Effects

In contrast to effects on IQ, effects on achievement do not fade out. In many studies, effects on achievement *appear* to fade out, but this is primarily due to substantial and selective attrition in follow-up that reduces the statistical power to detect effects and biases estimates of effects toward zero. Lasting effects on achievement are found in both experimental and quasi-experimental studies that do not suffer from these or other serious methodological flaws. True experiments with adequate sample sizes and minimal attrition find sizable effects on achievement test scores into adolescence.

Effects on School Success

There is highly uniform evidence of long-term positive effects on school success as measured by rates of grade retention, special education, and high school graduation. All but one of the model program studies reported grade retention and special education rates, and in each the rates are lower for the children with preschool education. Despite their small sample sizes, a statistically significant effect on grade repetition or special education rates was found in 5 model program studies and in one other for length of time in special education. Ten of 13 Head Start and public school program studies that collected relevant data reported statistically significant effects on grade retention or special education. In the 5 studies (3 model program, 1 Head Start, 1 public school) with high school graduation data, results favored the preschool education group in all 5 and were statistically significant in 3.

Comparing Model and Public Programs

From a public policy perspective, it is important to know how Head Start and public school early childhood programs compare to model programs with respect to effectiveness. Although both types of programs have been found to produce positive effects, the size of effects produced by the two types of programs may not be the same. Useful effect size comparisons are precluded for IQ and achievement test effects by the lack of IQ measures in Head Start and public school studies and the serious problems with achievement test data in many studies. However, it is possible to compare effect sizes for school success.

Average effects on cumulative rates of special education placement, grade repetition, and high school graduation by program type are presented in Table 1 [6]. Average effects are substantial for both types of programs, but effects on special education are much larger for model programs—24 percentage points for model programs and 5 percentage points for Head Start and public school programs. There is no statistically significant difference between the two types of programs in effects on grade retention. Effects are comparable for high school graduation, but the small number of studies with this outcome measure reduces confidence that the lack of difference in effects is representative.

Differences in effects on special education and grade repetition between the two program types were investigated further with regression analyses that controlled for design (randomized or not), length of follow-up, age of entry (prior to age 3 or not), type of program (model or public), and the comparison group's rate of grade retention or special education placement. With one exception, none of the independent variables was significantly related to the long-term effects. The comparison group's rate was positively related ($P < 0.01$) to the size of the estimated program effect. The higher the comparison group's rate, the larger a program's effect.

Although the regression results seem to indicate that Head Start and public school programs are as effective as model programs when the population served is taken into account, such a conclusion should be approached cautiously. Few studies of the two types of programs overlap in the degree of disadvantage indicated by the comparison group rates of special education or grade repetition. This raises questions about whether model programs have targeted more disadvantaged populations or public program studies have serious measurement problems. Also, program quality and intensity were at best crudely represented in the regression analysis. If public programs targeted more disadvantaged populations, they might be less successful unless they increased the quality and intensity of their services. Moreover, studies of program quality and one study directly comparing model and large-scale public programs indicate that differences in quality between model and public programs affect program outcomes [6].

TABLE 1

Long-Term Effects on School Success by Type of Program

Outcome measure	Model programs				Head Start/Public school			
	Median	Mean	SD	*N*	Median	Mean	SD	*N*
Decrease in special education	24.0	19.6*	14.6	11	5.0	4.7*	5.3	9
Decrease in grade retention	14.5	14.5	10.0	14	10.0	9.5	6.4	11
Increase in high school graduation	16.0	16.0	2.0	3	15.0	15.0	2.8	2

Note. Data were measured as percentage point decrease or increase.
*$P < 0.01$, *t* test.

ECONOMIC CONSEQUENCES

Programs that produce substantial improvements in the cognitive development and school success of children in poverty can be expected to produce substantial direct benefits through educational cost-savings and substantial indirect benefits as the result of increased productivity and social responsibility. A comprehensive cost–benefit analysis of the long-term effects of early childhood education has been conducted based on data from the High/Scope Perry Preschool study [7]. This study is a randomized trial of a part-day preschool education program with weekly home visiting that has collected detailed data on 123 study participants with minimal attrition through age 27 [8].

The results of the cost–benefit analysis are summarized in Table 2. All figures in the table are in 1992 dollars discounted at a real (i.e., inflation-adjusted) rate of 3%. The cost of the program was roughly $7600 per child for a year. The cost figure in Table 2 is a weighted-average of the costs of 1 and 2 years because a few children ($n = 13$) began at age 4 and received only 1 year of the program while most ($n = 45$) began at age 3 and received 2 years. Benefits were estimated in seven categories: custodial child care value, reduced costs of K–12 education, reduced costs of adult education, increased costs of college education, increased earnings and fringe benefits, decreased costs of crime, and decreased costs of welfare. These costs and benefits are shown for society as a whole and are broken down into the costs or benefits of direct effects on the study participants and the program costs and benefits of indirect effects on other members of society who may be thought of as the taxpaying public generally. For example, the economic benefit to society as a whole from a reduction in welfare dependency is merely the reduction in administrative costs, which is shown in the first column; program participants receive fewer welfare payments so this is shown as a cost in the second column, while the taxpayers benefit from both lower administrative costs and fewer payments so that the sum of these appears as a benefit in the third column. In addition, benefit estimates that depend only on program effects measured through age 27 are distinguished from benefit estimates that require projections beyond age 27 (such as increased earnings from ages 28–65 and future crime reductions).

As can be seen from the bottom lines of Table 2, the estimated economic benefits of preschool education are quite large relative to its costs. In fact, the estimated rate of return on preschool education exceeds the average rate of return on investments in the stock market over the past 30 years [7]. The generalizability of these results is enhanced by the fact that the Perry Preschool program's underlying effects are fairly close to the public program averages in Table 1 (special education 13 percentage points, grade repetition 5 percentage points, and high school graduation 18 percentage points) and

TABLE 2

Present Value of the Perry Preschool Program's Costs and Benefits per Child (1992 Dollars, Discounted at 3%)

Cost or benefit	A To society as a whole ($)	B Program participants ($)	C Taxpayers/ general public ($)
Measured benefits			
Child care	738	738	0
K-12 education	6,872	0	6,872
Adult education	283	0	283
College education	(−868)	0	(−868)
Earnings	14,498	10,270	4,228
Crime	49,044	0	49,044
Welfare	219	(−2,193)	2,412
Total measured	70,876	8,815	61,972
Projected benefits			
Earnings	15,833	11,215	4,618
Crime	21,337	0	21,337
Welfare	46	(−460)	506
Total projected	37,216	10,755	26,461
Total benefits	108,002	19,570	88,433
Program cost	(−12,356)	0	(−12,356)
Net benefits	95,646	19,570	76,077

Note. Costs or disbenefits appear as negative numbers in parentheses. Benefits reported under earnings include all employee costs paid by an employer including fringe benefits. Column A is the sum of columns B and C, but numbers may not add exactly to totals due to rounding. Source is Ref. [7].

by the magnitude of the economic benefits. Also, as Head Start and public school programs cost significantly less than the Perry Preschool program, they would pay off in the long-term even if their benefits were lower by a factor of 10.

PUBLIC POLICY IMPLICATIONS

In light of the evidence, every child living in poverty in the United States ought to be provided with at least 1 year of quality education prior to school entry in a part-day preschool education program or a full-day developmental child care program rich in cognitive interactions between teachers and children. Custodial child care, home-visiting programs, and other models that do not provide sustained, intensive improvements in the child's learning environment cannot be expected to produce the desired outcomes [3]. As there is some uncertainty about the size of the effects of existing programs, a conservative strategy would be to increase the quality and intensity of public programs to approach the levels of model programs found to be effective in the context of experimental evaluations of public programs' long-term effects.

Increased maternal labor force participation and federal welfare reform have made the half-day school-year preschool program at ages 3 and 4 obsolete for much of the population. As larger effects on cognitive development may be produced by full-day, year-round programs beginning before age 1, the government should sponsor large-scale experiments comparing the effects of such interventions to existing programs that begin at age 3 or 4. Head Start can conduct such randomized trials as well as investigating the potential returns to improving the quality of existing models. Although early childhood education is not a panacea, research-based early education programs can substantially improve the cognitive development, academic success, and lives of children in poverty while benefiting the nation as a whole.

REFERENCES

1. McKey RH, Condelli L, Ganson J, Barrett G, McConkey C, Platz M. The impact of Head Start on children, families, and communities. Washington: U.S. Govt. Printing Office, 1985. [DHHS Publication No. OHDS 85-31193.
2. White K, Casto G. An integrative review of early intervention efficacy studies with at-risk children: implications for the handicapped. Anal Intervent Dev Disability 1985;5:7–31.
3. Barnett WS. Long-term effects of early childhood programs on cognitive and school outcomes. Future Children 1996;5(3):25–50.
4. Caughy MO, DiPietro J, Strobino M. Day-care participation as a protective factor in the cognitive development of low-income children. Child Dev 1994;65,457–71.
5. Haskins R. Beyond metaphor: the efficacy of early childhood education. Am Psych 1989;44,274–82.
6. Barnett WS. Persistent effects on cognitive development and school success. In Barnett WS, Boocock SS, editors. Early care and education for children in poverty: promises, programs, and long-term results. White Plains: SUNY Press. In press.
7. Barnett WS. Lives in the balance: age-27 benefit–cost analysis of the High/Scope Perry Preschool program. Monographs of the High/Scope Educational Research Foundation (No. 11). Ypsilanti, MI: High/Scope Press, 1996.
8. Schweinhart LJ, Barnes HV, Weikart DP, Barnett WS, Epstein AS. Significant benefits: the High/Scope Perry Preschool study through age 27. Monographs of the High/Scope Educational Research Foundation (No. 10). Ypsilanti, MI: High/Scope Press, 1993.

EXECUTIVE SUMMARY

The High/Scope Perry Preschool Study Through Age 27

The High/Scope Perry Preschool Project is a study assessing whether high-quality, active learning preschool programs can provide both short- and long-term benefits to children living in poverty and at high risk of failing in school. For almost three decades, the study has followed the lives of 123 such children from African-American families who lived in the neighborhood of the Perry Elementary School in Ypsilanti, Michigan, in the 1960s. At the study's outset, the youngsters were randomly divided into a *program group*, who received a high-quality, active learning preschool program, and a *no-program group*, who received no preschool program. Researchers then assessed the status of the two groups annually from ages 3 to 11, at ages 14–15, at age 19, and most recently, at age 27, on variables representing certain characteristics, abilities, attitudes, and types of performance. The median percentage of missing cases for these various assessments was only 4.9%, and only 4.9% of cases were missing for the age-27 interviews. The study's design characteristics give it a high degree of internal validity, providing scientific confidence that postprogram group-differences in performance and attitudes are actually effects of the preschool program.

As shown in Figure 1, study findings at age 27 indicate that in comparison with the no-program group, the program group had

- Significantly[1] higher monthly earnings at age 27 (with 29% vs. 7% earning $2,000 or more per month)
- Significantly higher percentages of home ownership (36% vs. 13%) and second-car ownership (30% vs. 13%)
- A significantly higher level of schooling completed (with 71% vs. 54% completing 12th grade or higher)
- A significantly lower percentage receiving social services at some time in the previous 10 years (59% vs. 80%)
- Significantly fewer arrests by age 27 (with 7% vs. 35% having 5 or more arrests), including significantly fewer arrests for crimes of drug making or dealing (7% vs. 25%)

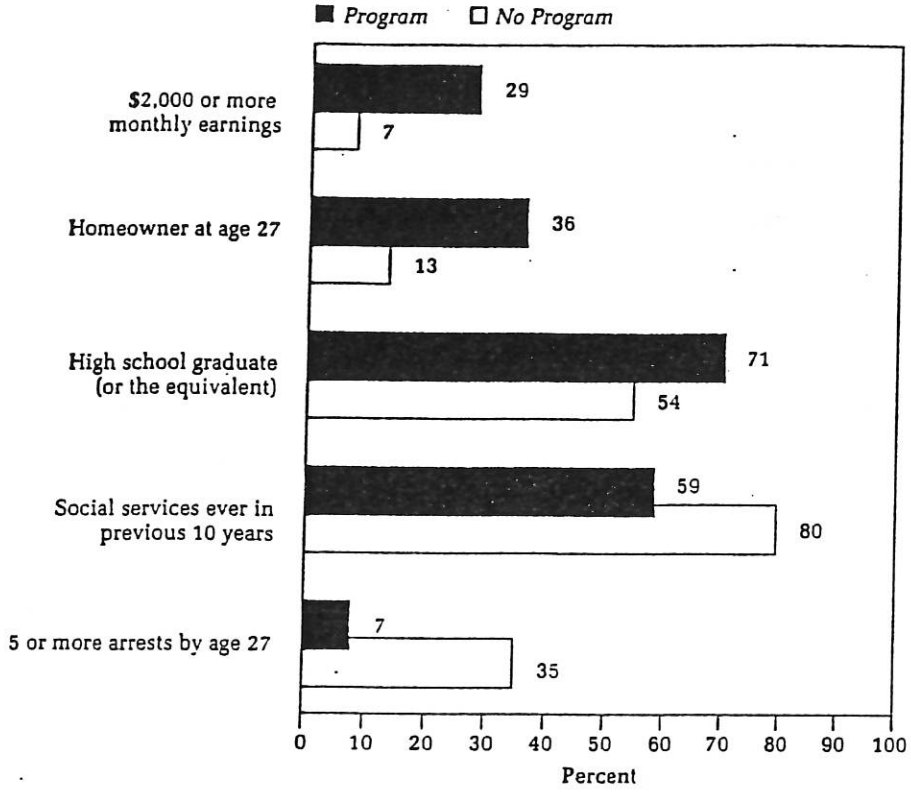
As shown in Figure 2, over the years the program group had significantly higher scores than the no-program group in

[1]This report describes a group difference as significant if it has a two-tailed probability of less than .05.

Schweinhart, L. J., H.V. Barnes, and D.P. Weikart. (1993). Significant Benefits: The High/Scope Perry Preschool Study Through Age 27. Monographs of the High/Scope Educational Research Foundation, No. 10, Ypsilanti, MI: High/Scope Press.

Figure 1

HIGH/SCOPE PERRY PRESCHOOL PROJECT — MAJOR FINDINGS AT AGE 27



Note. All findings are significant at $p < .05$, two-tailed. See Tables 9, 18, 22, 25, and 26 for details.

- General literacy at age 19 (on the Adult Performance Level Survey; American College Testing Program, 1976)
- School achievement at age 14 (on the California Achievement Tests; Tiegs & Clark, 1971)
- Intellectual performance (IQ) from the end of the first year of the preschool program to the end of first grade at age 7 (on the Stanford-Binet Intelligence Scale, Terman & Merrill, 1960)

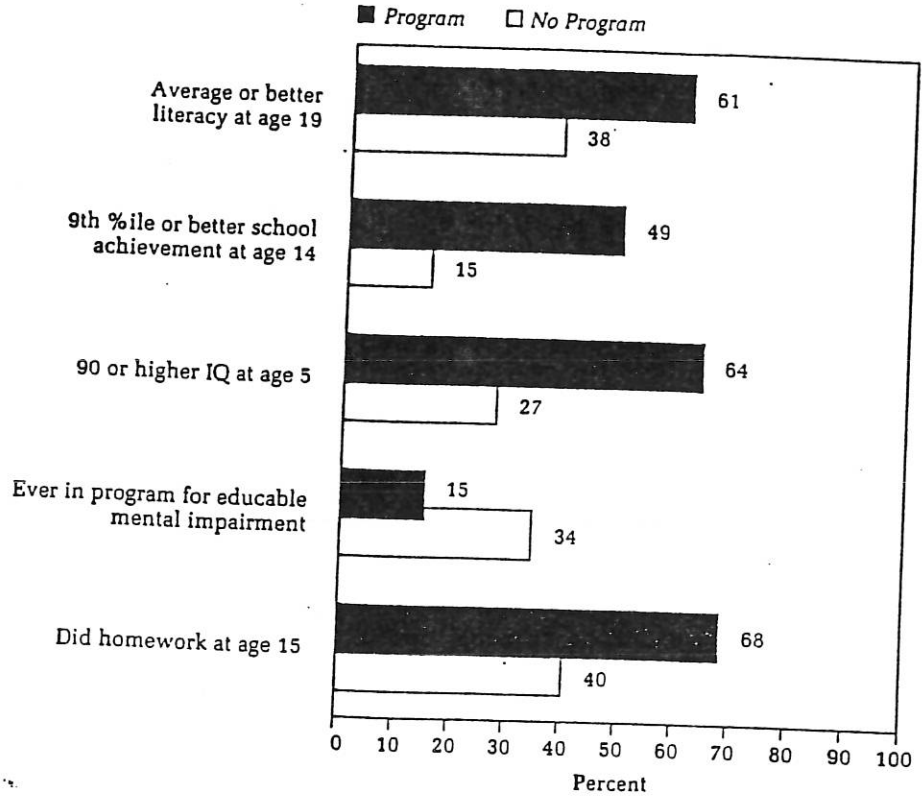
Moreover, as compared with the no-program group, the program group

- Spent significantly fewer years in programs for educable mental impairment (with 15% vs. 34% spending a year or more in EMI programs)
- Had a significantly higher percentage reporting at age 15 that their school work required preparation at home (68% vs. 40%)

As a group, the program females reported significantly higher monthly earnings at age 27 than the no-program females did (with 48% vs. 18% earning over $1,000) because more of the program females (80%

Figure 2

HIGH/SCOPE PERRY PRESCHOOL PROJECT —
MAJOR EDUCATIONAL PERFORMANCE FINDINGS



Note. All findings are significant at $p < .05$, two-tailed.
See Tables 10, 12, 13, 14, and 15 for details.

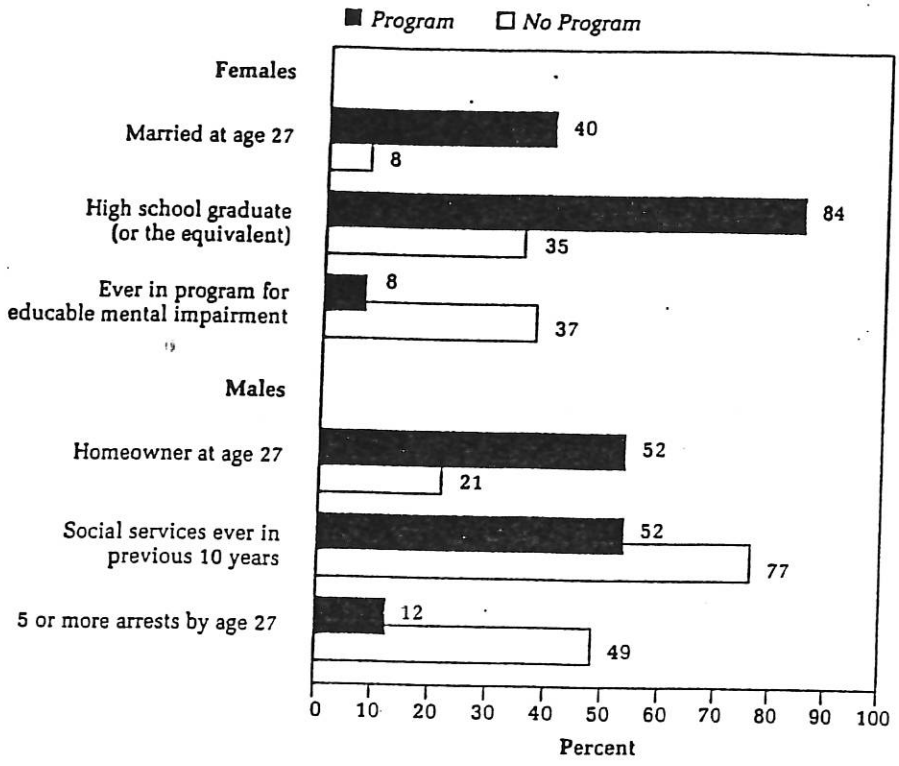
vs. 55%) had found jobs. The program males, as a group, had significantly higher monthly earnings at age 27 than the no-program males (with 42% vs. 6% earning over $2,000) because the program males had better paying jobs. Of employed males in the two groups, 53% vs. 8%, respectively, were earning over $2,000, which is a significant difference.

As shown in Figure 3, unlike the significant differences in monthly earnings, which were found for *both* females and males, the following significant differences between the program group and the no-program group at age 27 were found to hold for males only or for females only. As compared with no-program females,

- Significantly more program females were married at age 27 (40% vs. 8%), and significantly fewer of the births to program females were out of wedlock (57% vs: 83% of births).
- Significantly more program females completed 12th grade or higher (84% vs. 35%).
- Significantly fewer program females spent time in programs for educable mental impairment (8% vs. 37%).

Figure 3

HIGH/SCOPE PERRY PRESCHOOL PROJECT —
MAJOR FINDINGS THROUGH AGE 27, BY GENDER



Note. All findings are significant at $p < .05$, two-tailed.
See Tables 11, 19, 25, and 28 for details.

As compared with no-program males,

- Significantly more program males owned their homes at age 27 (52% vs. 21%).
- Significantly fewer program males received social services at some time between ages 18 and 27 (52% vs. 77%).
- Significantly fewer program males had 5 or more lifetime arrests (12% vs. 49%).

The findings listed here have economic values that are benefits to society. Compared with the preschool program's cost, these benefits make the program a worthwhile investment for taxpayers as well as for society in general. **Over the lifetimes of the participants, the preschool program returns to the public an estimated $7.16 for every dollar invested.**

These findings are remarkable. Their positive implications for improved quality of life for participants, their families, and the commu-

nity at large are of tremendous importance. No evidence available when this study was initiated even suggested that a preschool program for children 3 and 4 years old could set in motion a chain of events leading to such lifetime effects on the children. Yet the internal validity of the study constitutes a powerful argument that all the outcomes stated here are in fact due to the program that the young children experienced at the outset of the High/Scope Perry Preschool Project.

When the preschool studies of the 1960s were begun, people spoke of "breaking the cycle of poverty" and "inoculating" children against failure. However, life and poverty are much more complicated than this all-or-nothing rhetoric suggests. This study nevertheless shows that amidst this complexity, a good program can make significant differences. The rhetoric of the 1990s should reflect this more sophisticated understanding of the complexity of poverty. *People don't break the cycle of poverty: Some people ease out of it a little. The preschool "inoculation" is not universally successful, like polio or smallpox vaccinations. But like many other medicines, it brings about significant differences.*

The High/Scope Perry Preschool Project findings present us with the challenge to develop and maintain widespread preschool programs similar to the program developed for this study. Such a program, like other preschool programs that have demonstrated lasting success, would

- Provide for all 3- and 4-year-olds living in poverty a classroom program operating at least 12½ hours each week
- Use developmentally appropriate practices that encourage child-initiated learning activities
- Have a high level of outreach to parents as partners
- Maintain a child-staff ratio of no more than 10 children per adult
- Employ staff who are well trained in early childhood education
- Provide consistent staff supervision and staff training in the developmentally appropriate curriculum approach used

It is essential that we invest fully in high-quality, active learning preschool programs for all children living in poverty. Since the national Head Start program and state-funded preschool programs now serve fewer than half of these most vulnerable of our children, the nation is ignoring tremendous human and financial potential.

- We must spend what it takes. This means full funding for the national Head Start program and similar programs—enough to not only *serve all eligible children but also adequately serve each child,* so programs can help all children reach their potential.
- We must ensure that the policies and procedures in place for all early childhood programs are of the quality necessary to permit staff to do their jobs well.

- Early childhood program staff must do their jobs well, bringing to them a firm sense of purpose, a deep love of children, and solid competence in early childhood education.

Doing all this is sure to bring significant benefits to the lives of the next generation.

Evaluation of a Full-Day Kindergarten Program

Y. Lawrence Wang and Whitcomb G. Johnstone

Many school districts are moving toward full-day kindergarten to accommodate the needs of working families; others maintain half-day kindergartens, or offer families a choice. When deciding what kind of kindergarten program to offer, educators must consider not only the desires of parents but also the potential effects on student achievement.

In this article, authors Y. Lawrence Wang and Whitcomb G. Johnstone investigate differences in oral language development, emergent literacy skills, mathematical reasoning and concepts, and behavior for full-day vs. half-day kindergarten students in the Irving (Texas) Independent School District. The study found some statistically significant advantages for the full-day kindergarten program, and supported the school district's decision to expand the full-day program.

In school year 1995-96, the Irving Independent School District piloted a full-day kindergarten program. As stated by Dr. Martha Stone, assistant superintendent for curriculum and instruction, two of the major objectives for full-day kindergarten are to:

1) help all kindergarten students reach a level of oral language development and literacy sufficient to be successful in first grade, and

2) decrease or eliminate the use of pre-first grade or any form of grade retention with students at the primary level.

In the long term, the district expects full-day kindergarten, in combination with other early intervention strategies, to contribute to stronger performance beyond the primary grade levels on state and national assessments.

Each school designed its own full-day kindergarten program. However, all of the programs subscribed to the following common objectives:

- improved oral language development;
- improved "emergent" reading skills;
- enhanced acquisition of early mathematics concepts; and
- improved appropriate behaviors.

This article discusses results of the second-year evaluation of the full-day kindergarten program, which was undertaken to help determine whether the district should expand the full-day program to more elementary schools. A major argument for full-day kindergarten is that additional hours in school would better prepare children for first grade and would result in a decreased need for grade retention, including pre-first, and fewer referrals to special education. Therefore, in this evaluation study, our research question is whether, given more hours of instruction, students in the full-day kindergarten program exhibit greater growth in oral language development, reading skills, mathematics concepst, and appropriate behaviors than students in the half-day program.

Design and Method

Participants –

This evaluation consisted of four parts. Each part investigated one of the four objectives of the full-day kindergarten program described above. A stratified

When this article was written, Y. Lawrence Wang was Measurement and Evaluation Specialist for the Department of Planning, Evaluation and Research of Irving Independent School District, Texas (enrollment 26,500). Whitcomb G. Johnstone is Director of the Department of Planning, Evaluation, and Research of Irving Independent School District.

1-43

Table 1.—The Number of Classes and Students Involved in Each Sub-Study

Program	Math Proficiency		Emergent Literacy		IPT Oral Language		Social Development Checklist	
	Classes	Students*	Classes	Students*	Classes	Students*	Classes	Students*
Full-day K (Regular)	23	412	22	415	23	308	15	198
Half-day K (Regular)	12	143	11	135	17	148	6	51
Full-day K (Bilingual)	5	80	5	84	7	78	3	10
Half-day K (Bilingual)	7	45	7	101	6	30	3	14

*tested in both fall 1996 and spring 1997

random sampling strategy was applied so that all samples were drawn proportionately from each school for each analysis. This sampling strategy was used to balance the impacts of such factors as school effects or geographical (school cluster) effects.

All kindergartners who stayed in the same program throughout the pre- and post-assessment period were included in at least one of the four sub-studies. About half of the kindergartens in the district participated in the emergent literacy study, and the other half (45 classes) participated in the mathematics study. See Table 1 for the number of classes and students in each study.

Instruments—

Oral Language Development. We selected the IPT Oral Language Assessment (Ballard, Tighe, and Dalton 1991) as a pre- and post-measure of the students' oral language development. Because we wanted to assess the development of the language of instruction, students in the monolingual English classes were tested with the English version of the IPT and students in the bilingual classes were tested with the Spanish version (Ballard, Tighe, and Dalton 1989). Our experience with the IPT Oral Language Assessment in our first-year evaluation of the full-day kindergarten pilot demonstrated that raw score gains provide a more sensitive measure of development than change in language level.

Mathematics Concepts. The concepts and reasoning section of the Woodcock-McGrew Werder Mini-Battery of Achievement (MBA) (Woodcock, McGrew, and Werder 1994) was used as the measure of mathematics development. Because there was no Spanish MBA, the first 35 items were translated into Spanish locally for use with the bilingual group, with permission from the test publisher.

Literacy Skills. "An Observation Survey of Early Literacy Achievement" (Clay 1995) was used to assess emergent literacy in five skill areas: letter identification; word recognition; concepts of print; writing vocabulary; and recognizing sound in words. Students in the English group were tested using the English version of this instrument (Clay 1995), and bilingual students were tested using the Spanish version (Clay, et al. 1996).

Behavior. For this study, the social/emotional development checklist on the district's new kindergarten report card was used. There are eight facets of student behavior in the checklist: follows classroom/school rules, accepts responsibility for own actions, works well independently, works and plays well with others, uses time wisely, uses self-discipline, cares for property and materials, and participates in class activities.

At the end of each six-week period, teachers mark an "X" at the areas in which they feel improvement is needed. The sum of the number of marks for the first three six-week periods was computed as the first-semester score, and the marks for the remaining three six-week periods were combined as the second semester score. All scores were then translated into a four-point scale ranging from 0 to 3. In this way, students who received no check marks in any of the three six-week periods would receive three points, indicating the most positive level of at-school behavior. On the other hand, a "0" score would indicate the most negative level of behavior, and would be assigned to students who received check marks for all three reporting periods.

Pre- and Post-Assessment—

Teachers who were in the reading or the mathematics sub-study were instructed to complete the pre-assessments with all students in their designated classes by the end of October 1996. Post-assessments were completed by the next-to-last week of the school year in May 1997. A one-day substitute teacher was provided to assist the kindergarten teachers with the children as they completed these assessment in both semesters.

1-44

Under the coordination of Barbara Neal and Ron Robertson at the bilingual/ESL center, the pre-assessment of IPT Oral Language was accomplished by campus personnel with the assistance of bilingual/ESL assessment center staff during the second week of October. The post-assessment was accomplished in the same way about one week prior to the end of the school year.

Three weeks before the end of the school year, teachers in the behavior sample were instructed to send copies of the student report cards for all students in their classes to the planning, evaluation, and research office. Most of the report cards arrived at the office of planning, evaluation, and research by the end of the second-to-last week of the school year.

Findings

This study focused on one primary research question: "Does full-day kindergarten enhance students' oral language development, emergent reading skills, early mathematics reasoning, and appropriate behaviors to a greater level than a half-day program?" Our hypothesis, assuming similar curriculum and instruction, was that students in full-day kindergarten, with a longer instructional day, would show greater gains than students in half-day kindergarten on the IPT Oral Assessment, the emergent literacy observation forms, the MBA mathematics concepts and reasoning assessment, and the social emotional development checklist in the report card.

Oral Language Development –

During the first six-week and the last six-week periods, IPT Oral Language Assessments were administered to assess students' oral language development. The total number of correct items was recorded at each assessment. A statistical procedure known as the repeated measures analysis of variance was performed on the scores. Only scores for the 564 students who stayed in the same program throughout the pre- and post-assessment period were included in the analysis. Results of the English-speaking and Spanish-speaking classes were analyzed separately.

Table 2 shows the mean pre-test and post-test scores on the IPT for English and Spanish-language students in each program districtwide; Figure 1 on page 30 shows gains for these groups from pre-test to post-test. Students in full-day group started and ended higher on the IPT than students in the half-day group. Both full-day and half-day students made progress in oral language development, but the full-day kindergarten students made greater gains, for both the English-speaking and the Spanish-speaking groups.

Mathematics Concepts and Reasoning –

The Concepts and Reasoning Section of the mathematics test in the Mini-Battery of Achievement (MBA) was used to measure students' early mathematics concepts. This section consists of 50 items, distributed uniformly in respect to difficulty. A Spanish translation of the first 35 items was done locally with permission from the publisher.

Table 2.—Pre-test and Post-test Mean Raw Scores on the IPT Oral Language Assessment

	English		Spanish	
	Half-day K (n=148)	Full-day K (n=308)	Half-day K (n=30)	Full-day K n=78)
Pre-test	34.04	35.94	18.10	21.22
Post-test	47.68	51.91	25.80	33.10
Gains	13.65	15.97	7.70	11.88

Note: IPT Score Range= 0-83

Table 3 lists the mean pre-test and post-test raw scores on the MBA Concepts and Reasoning Test for English and Spanish groups. The analyses of the mean gains in mathematics concepts and reasoning for these groups are shown in Figure 2 on page 30.

Table 3.—Pre-test and Post-test Mean Raw Scores on the MBA Mathematics Reasoning and Concepts Test

	English		Spanish	
	Half-day K (n=143)	Full-day K (n=412)	Half-day K (n=45)	Full-day K (n=80)
Pre-test	17.17	19.25	17.71	18.21
Post-test	22.48	25.36	23.56	23.84
Gains	5.33	6.11	5.85	5.63

Note: MBA Score Range=0-50; 0-35 in Spanish

Emergent Literacy Skills –

Each of the five literacy observation surveys targets one particular facet of a child's emergent literacy skills: letter identification, word recognition, concepts of print, writing vocabulary, and recognizing sounds in words. To examine the kindergarten program's effect on overall literacy skills, a statistical technique known as factor analysis was applied to the five literacy observation scores. The results of the factor analysis indicated that these five literacy observation measures seemed to measure one general factor. On the basis of this finding, we created a composite "literacy" score based on the percentage of correct responses to all five literacy surveys and used it as an indicator of overall literacy growth.

1-45

Figure 1.—Pre-test and Post-test Gains on Kindergarten IPT Oral Language Assessment

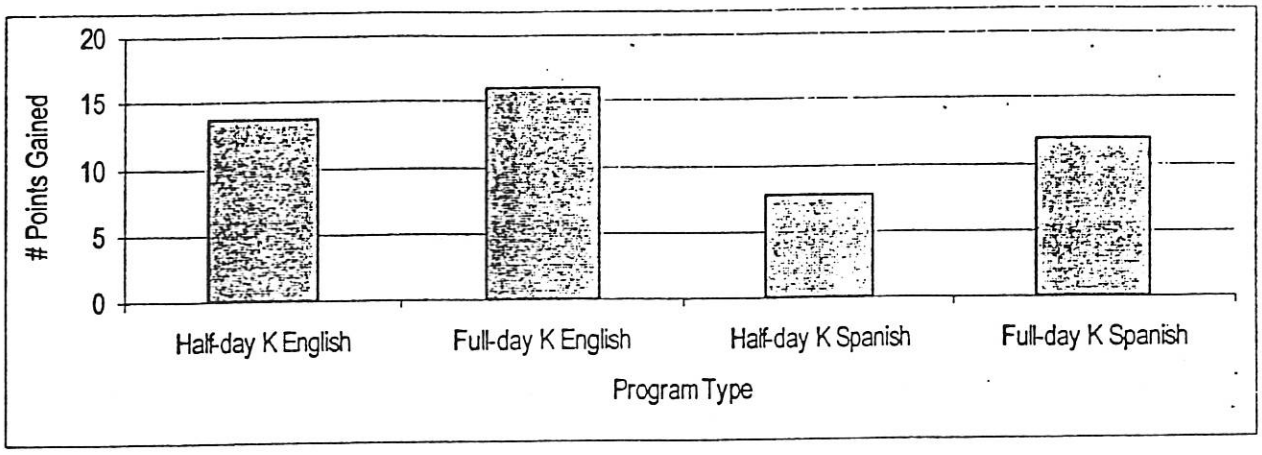


Figure 2.—Pre-test and Post-test Mean Raw Score Gains on the MBA Mathematics Reasoning and Concepts Test (MBA Score Range=0-50; 0-35 in Spanish)

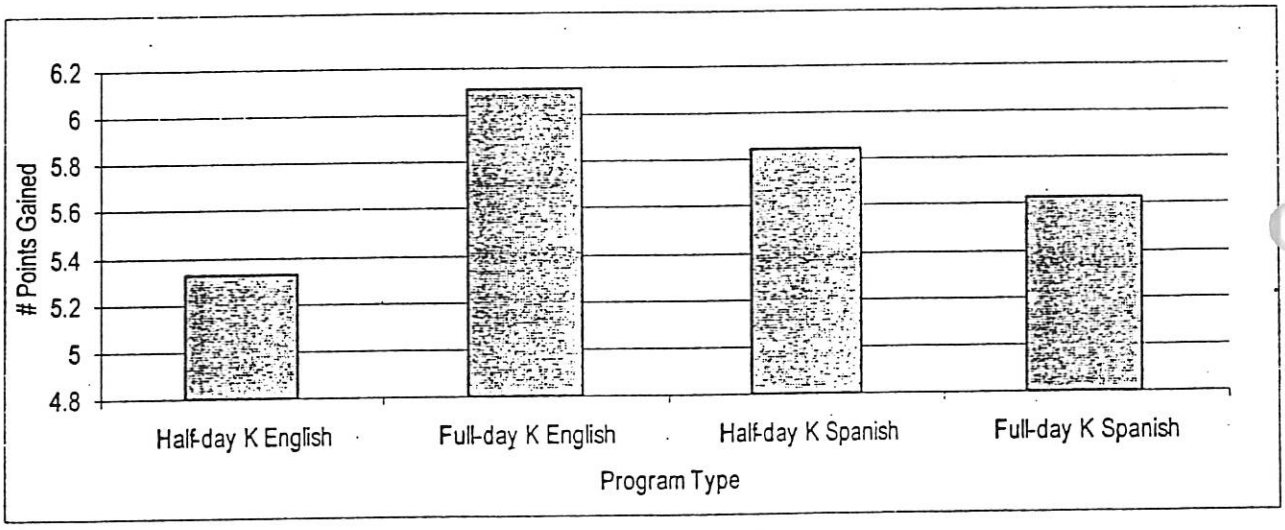
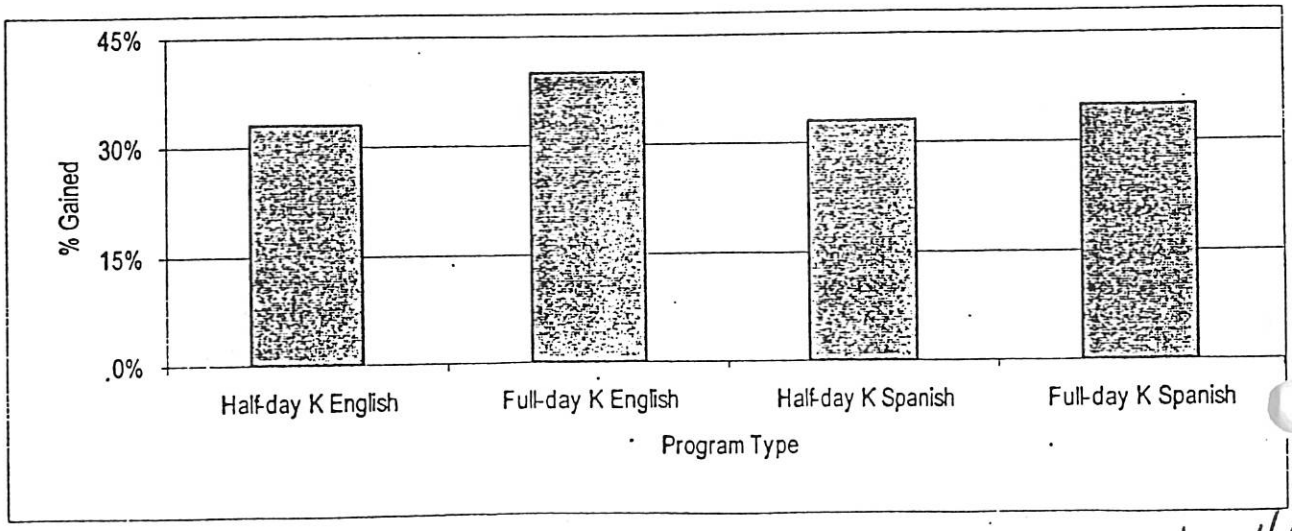


Figure 3.—Percent of Correct Item Response Gains on the Observation Survey of Early Literacy Achievement (Composite)



1-46

Table 4.—Percent of Correct Item Responses on the Observation Survey of Early Literacy Achievement (Composite)

	English		Spanish	
	Half-day K (n=121)	Full-day K (n=378)	Half-day K (n=61)	Full-day K n=81)
Pre-test	31%	30%	12%	14%
Post-test	64%	70%	45%	49%
Gains	33%	40%	33%	35%

Table 5.—Mean Scores on the Irving ISD Report Card Social/Emotional Developmental Checklist (Composite)

	English		Spanish	
	Half-day K (n=51)	Full-day K (n=198)	Half-day K (n=14)	Full-day K (n=10)
Pre-test	21.45	21.05	15.21	7.30
Post-test	21.55	22.09	18.29	11.20
Gains	0.10	1.04	3.08	3.90

Table 4 above displays the percentage of correct responses for half- and full-day English-speaking and bilingual classes. The mean gains on the composite emergent literacy scores for English-speaking and Spanish-speaking groups are displayed respectively in Figure 3 on page 30.

Figure 3 shows significant full-day versus half-day gains on the composite literacy score in favor of the full-day program in the English-speaking classes. However, there was no statistically significant difference in the Spanish-speaking full-day versus half-day comparison.

Social/Emotional Development –

The social and emotional checklist on the kindergarten student report card, which covers eight areas of school behavior, was used as a measure for student behaviors in this study. Our research question was whether the teachers in the full-day classes would rate their students as displaying fewer inappropriate behaviors than teachers in the half-day program.

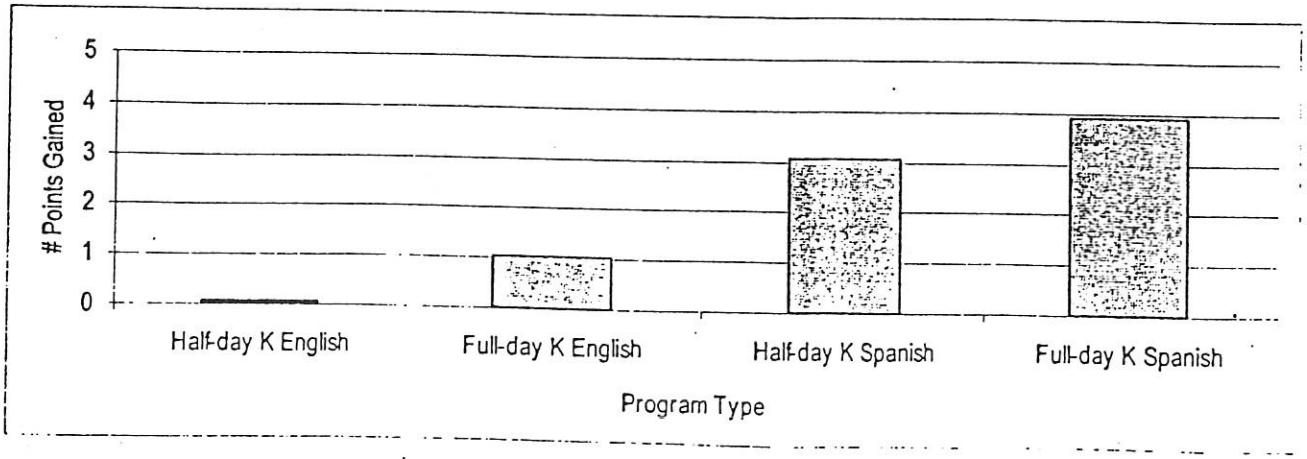
The results of a factor analysis similar to the previous one for the literacy measures indicated that using one overall summary score for the eight social-emotional indicators was appropriate. These indicators were all on a 0 to 3-point scale in which a higher score represented better behavior. Summing the scores for all eight indicators would result in a possible highest score of 24. The average scores for both kindergarten groups on this overall behavior scale are displayed below in Table 5.

The gains for English-speaking and Spanish-speaking groups are shown in Figure 4 below, indicating that full-day students in both English-speaking and Spanish-speaking classes made greater gains than their half-day peers. No statistical analysis was applied to the behavior scores due to extreme skew in their distribution.

Summary and Conclusions

The findings of this second-year district-wide evaluation study of full- and half-day kindergarten programs basically coincide with the findings in the first-year pilot. The results of statistical analyses conducted for the English-speaking full-day versus half-day comparison in oral language development, early mathematics concepts, and emergent literacy skills all indicated significant differences in favor of full-day kindergarten. The analysis of pre- to post-test

Figure 4.—Mean Scores Gains on the Irving ISD Report Card Social/Emotional Development Checklist (Composite)



1-47

ratings of school behavior also indicated that being in a full-day program tended to improve students' behaviors more than being in a half-day program.

The results of the same full- vs. half-day analyses for Spanish-speaking classes did not find as many statistically significant differences between these two groups of students. However, within the Spanish classes the analyses of the pre- to post-test differences still indicated that full-day students tended to make more improvement than half-day students in oral language, mathematics concepts, and emergent literacy.

The lack of statistical significance in the gains for the Spanish-speaking group may be a reflection of the smaller number of participants in that group compared to the English-speaking group. The pattern of gains between full- and half-day programs were similar for both groups.

Overall, the weight of the findings favor the full-day kindergarten program and support the decision to expand full-day kindergarten to all elementary schools in the Irving school district. ❑

References

Ballard, W. S., P. L. Tighe, and E. D. Dalton. 1991. *IDEA Oral Language Proficiency Test (Form C)*. Brea, CA: Ballard and Tighe, Inc.

Ballard, W. S., P. L. Tighe, and E. D. Dalton. 1989. *IDEA Oral Language Proficiency Test (Spanish Version)*. Brea, CA: Ballard and Tighe, Inc.

Clay, M. M. (1995) *An Observatiuon Survey of Early Literacy Achievement* (3rd ed.) Portsmouth, NH: Heinemann Education.

Clay, M. M., K. Escamilla, M. A. Andrade, A. G. Basurto, and O. A. Ruiz. 1996. *Spanish Reconstruction of an Observation Survey (A Bilingual Text)*. Portsmouth, NH: Heinemann Education.

Woodcock, R. W., K. S. McGrew, and J. K. Werder. 1994. *Woodcock-McGrew-Werder Mini-Battery of Acheivement*. Chicago, IL: Riverside Publishing.

1-48

DiPasqua W., Moule, A.D., & Flewelling, R.W. (1980). The birthdate effect. *Journal of Learning Disabilities, 13,* 234–237.
-t, B. (1973). Influence of teacher behavior in the preschool. *Developmental Psychology, 9,* 198–206.
eeman, E.B. (1990). Issues in kindergarten policy and practice. *Young Children, 45*(4), 29–34.
Goffin, S.G. (1989). Developing a research agenda for early childhood education: What can be learned from the research on teaching? *Early Childhood Research Quarterly, 4,* 187–204.
Gullo, D.F. (1990). The changing family context: Implications for the development of all-day kindergarten. *Young Children, 45*(4), 35–39.
Gullo, D.F. (1991). The effects of gender, at risk status and number of years of preschool on children's academic readiness. *Early Education and Development, 2,* 32–39.
Gullo, D.F., & Clements, D.H. (1984). Differences in achievement patterns for boys and girls in kindergarten and first grade: A longitudinal study. *Psychological Reports, 54,* 23–27.
Karweit, N. (1988). Quality and quantity of learning time in preprimary programs. *Elementary School Journal, 89.* 119–123.
Katz, L. (1988, Summer). What should young children be doing? *American Educator,* 28–33, 44–45.
Kinard, E.M., & Reinherz, H. (1986). Birthdate effects on school adjustment: A longitudinal study. *Journal of Educational Research, 79,* 366–372.
Lee, P.C. (1973). Male and female teachers in elementary schools. *Teacher College Record, 75,* 79–98.
Lehr, F. (1982). Cultural influences and sex differences in reading. *The Reading Teacher, 35,* 744–746.
Leinhardt, G., Sewald, A.M., & Engel, M. (1979). Learning what's taught: Sex differences in instruction. *Journal of Educational Psychology, 71,* 432–439.
Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.
Mitchell, A. (1989). Old baggage, new visions: Shaping policy for early childhood programs. *Phi Delta Kappan, 70*(9), 665–672.
National Association of State Boards of Education. (1988). *Right from the start.* Alexandria, VA: Author.
Nurss, J.R., & McGauvran, M.E. (1974). *Metropolitan Readiness Test (Level II, Form P).* New York: Harcourt Brace Jovanovich.
Powell, D.R. (1987). Comparing preschool curricula and practices: The state of research. In S. Kagan & E. Zigler (Eds.), *Early schooling: The national debate.* New Haven, CT: Yale University Press.
Sava, S.G. (1987). Development, not academics. *Young Children, 42*(5), 15.
Schweinhart, L.J., Weikart, D.P., & Larner, M.B. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly, 1,* 15–45.
Shepard, L.A., & Smith, M.L. (1986). Synthesis of research on school readiness and kindergarten retention. *Educational Leadership, 44*(3), 78–86.
Shepard, L.A., & Smith, M.L. (1988). Escalating academic demand in kindergarten: Counterproductive policies. *The Elementary School Journal, 89,* 135–145.
Uphoff, J., & Gilmore, J.E. (1986). *Summer children: Ready or not for school.* Middletown, OH: J & J Publishing.

Success Outcomes of Full-Day Kindergarten: More Positive Behavior and Increased Achievement in the Years After

John R. Cryan
University of Toledo
Robert Sheehan
Cleveland State University
Jane Wiechel
Irene G. Bandy-Hedden
Ohio Department of Education

This statewide longitudinal study was designed to investigate the effects of kindergarten schedule (half day, alternate day, and full day) and prior preschool attendance on elementary children's success (achievement, incidence of grade retention, provision of special educational services, and classroom behavior). Academic data are summarized from two phases of the study: a retrospective analysis of children's outcomes related to kindergarten attendance in 27 school districts in the years 1982, 1983, and 1984; and a prospective analysis of two cohorts of children, one entering kindergarten in fall 1986 in 27 school districts and one in fall 1987 in 32 school districts. Behavioral outcome data are reported in detail. Existing data found in cumulative folders, representing scores from 13 different standardized tests, and various outcome data were analyzed for the retrospective study. Outcome data for the ongoing study were gathered from the Metropolitan Readiness Test (administered in kindergarten), the Metropolitan Achievement Tests (administered in first grade), and the Hahnemann Elementary School Behavior Rating Scale (administered in kindergarten). Results from the longitudinal study indicate that children who attend preschool prior to kindergarten experience greater subsequent success in elementary school than those who do not. Results from both phases of the study indicate that participation in full-day kindergarten is positively

Correspondence and requests for reprints should be sent to John R. Cryan, University of Toledo, College of Education and Allied Professions, Department of Elementary and Early Childhood Education, Toledo, OH 43606-3390.

related to subsequent school performance, at least through first grade. Additional analyses demonstrate the significant impact of age at entrance to kindergarten and of gender.

Most 5-year-old children in the United States (86.5%) go to kindergarten (Karweit, 1988). The majority of those children have typically attended a traditional half-day program, although the numbers attending full-day kindergarten are increasing (Olsen & Zigler, 1989). The pressures of changing economic and social conditions in our society, however, are forcing educators and parents across the country to reconsider this "traditional" half-day kindergarten attendance pattern, replacing it with the all-day (or full-day) kindergarten. In a recent article, Gullo (1990) discusses how the all-day kindergarten seems better suited to respond not only to society's changing social and economic needs, but to children's educational needs as well. Walsh (1989) cautions that policy shifts and resultant changes in attendance patterns should be based upon more than needs. He calls for reliance upon good theory and solid evidence.

In 1985, the Ohio Department of Education (Department) was seeking research information to inform statewide policy-making in the area of early childhood education. Because policy decisions such as funding for public preschool and mandated all-day kindergarten were being considered, the Department sought conclusive information on the effects of various kindergarten schedules. Moreover, interest was expressed in the value of the preschool experience prior to kindergarten.

Exhaustive reviews of the literature on kindergarten schedules (Cryan, 1986; Helmich & Wasem, 1985; Karweit, 1987; Nurss & Hodges, 1982) document that existing studies of the effects of different kindergarten schedules (a) tend to be with small samples or unique populations, (b) generally fail to use rigorous research standards, (c) give almost exclusive priority to academic outcomes, and most important, (d) offer little or no convincing evidence favoring one type of schedule over another.

Prior attendance at preschool, however, is clearly linked to the successes disadvantaged children have in school. Any doubt about this controversial issue was laid to rest with the publication of the High/Scope research findings (Barnett & Escobar, 1987; Randolph, 1986; Schweinhart, Weikart, & Larner, 1986), and buttressed by the Cornell Consortium data (Darlington, Royce, Snipper, Murray, & Lazar, 1980) and recent findings of the Department of Health and Human Services (Copple, Cline, & Smith, 1987). Karweit (1989a, 1989b) cautions however, that effects of Head Start and ‘ier preschool programs often disappear within 2 to 3 years. Preschool is a potentially cost-effective method for reducing the negative effects of poverty xperienced by young children, but is it cost-effective for *all* children?

Given the enormous expense of a mandate to fund public preschool and full-day kindergarten, coupled with the general lack of data to support the latter approach, the Department initiated a 4-year statewide study. The specific purpose of the study was to examine the effects of kindergarten schedule and prior preschool attendance on children's academic and behavioral successes in kindergarten and beyond. These dependent (outcome) variables were thought to have equal weight among the members of the committee who planned the effort. That is, influenced by Elkind (1988) and others (NAEYC), and concerned that young children are being unnecessarily stressed by participating in extended school experiences focusing heavily on academics, we chose to question whether preschool and/or a full-day kindergarten experience could amount to more than academic success. Would children who went to preschool and/or a longer kindergarten behave differently? Only one study (Gullo, Bersani, Clements, & Bayless, 1986) looked at behavioral outcomes of different kindergarten schedules, and the results tended to mirror those of the studies focusing on academic outcomes. Once the initial planning for the study was complete, interest developed further about the additional independent variables of age at entrance to kindergarten and of gender.

The age at which a child can enter kindergarten is, for the most part, determined by state policy. The most frequent cutoff age is for a child to have turned 5 by September 1 of the year in which he or she enters kindergarten. Children in Ohio must have turned 5 by September 30 of the year in which they enter kindergarten.

The phenomenon of the "Summer Child," a term used to refer to those children who turn 5 in the summer before their entrance to kindergarten, has been discussed at length (Campbell, 1985; Diamond, 1983; Uphoff & Gilmore, 1985, 1986). All states permit such children to enter kindergarten, as no states have a June 1 cutoff date (Harris & Harris, 1986). There are actually two types of summer children, (a) those who attend kindergarten when they are age-eligible (youngest in their age group), and (b) those whose entrance to kindergarten is delayed by a full year after they are age-eligible (oldest in their age group). Our data indicate that in Ohio more than one third of the parents of summer children decide to delay their child's entrance to kindergarten. The studies cited above have helped foster a widespread belief that summer children (particularly boys) are at risk for school failure. Shepard and Smith (1986) contend that the "risk" is exaggerated and that there is only a 7 or 8 percentile-point difference separating the oldest from the youngest in first grade. They further state that this difference disappears by third grade.

Gender differences favoring females in verbal skills and males in spatial and mathematical skills are consistently supported in the research literature (Maccoby & Jacklin, 1974). Dwyer (1973) points out that the differences in

fe ...s unfold from the start of school. Male superiority in mathemati... do... ...t appear so early and may well be delayed until adolescence. Coates (1974) even suggests that girls may outperform boys in the preschool years. Plomin and Foch (1981) point out that, although there seem to be consistent differences between sexes, the magnitude of these differences is quite small. In the Maccoby and Jacklin (1974) data, for example, gender accounted for 1% to 2% of the reading performance variance and about 4% of the mathematics performance variance. Moreover, there is considerable variability between studies, suggesting the possibility that patterns of sex differences may differ for different populations.

This article seeks then to describe the entire complex of studies begun by the Ohio Department of Education in 1985. Because the academic outcome data are presented in depth elsewhere (Sheehan, Cryan, Wiechel, & Bandy-Hedden, 1991), this article provides only a brief summary of those effects. Following that, we describe in detail the children's behavioral outcomes assessed by teachers in relation to kindergarten schedule, preschool attendance, age at entrance to kindergarten, and gender.

METHOD

Design

The study was conceptualized in two phases. The first was a retrospective analysis of children who entered kindergarten in 27 school districts in the fall of 1982, 1983, and 1984. The second was a longitudinal study of children who entered kindergarten in 120 classes in 27 school districts in fall 1986 (Cohort 1) and 132 classes in 32 school districts in fall 1987 (Cohort 2).

School districts in the study were carefully chosen to provide comparisons of the effects of three kindergarten schedules:

1. Half day (typically 5 days per week, 2.5 hours per day)
2. Alternate day (typically 5 days in 2 weeks, 5 hours per day)
3. Full day (typically 5 days per week, 5 hours per day)

Whenever possible, school districts with two or more kindergarten schedules operating concurrently in the district (or in a building) were selected to participate in the study. When school districts offered only one kindergarten schedule, they were matched with a contiguous, demographically similar school district in the same county.

Defining Age

In defining the age factor, we categorized the kindergarten pupils with valid birthdates into one of five groups. The data is displayed in tabular form (see ...able 1). We note the number of children who are in Age Group 4. These

Table 1. Percent of Entering Kindergartners by Age Group

Group	Age October 1 of Kindergarten Year (In months)	Cohort 1 (%)	Cohort 2 (%)
1*	≤64	26	26
2	65–68	31	32
3	69–71	31	29
4**	72–75	9	9
5***	≥76	2	3
		99	99

* This group represents Summer Children who attended kindergarten as the youngest children in the class.

** This group represents Summer Children who could have attended kindergarten during the previous year but attended a year late.

*** This group represents children who entered kindergarten a full year and 3 months behind their classmates.

are children who could, based upon age, have entered kindergarten a full year earlier, but whose parents delayed their entrance until the following year. Anecdotally, we are aware of many parents making this decision for fear their summer children will experience school failure if they begin kindergarten as the youngest of their classmates. The decision by many parents to delay their kindergarten child's entrance a full year further increases the apparent immaturity of those kindergarten children who do attend as the youngest of their class.

Data Collection

Outcome data gathered on the 8,290 pupils in the retrospective study consisted of grade retention information, provision of special educational services, provision of remedial educational services, and whatever standardized test data had been routinely collected by school districts (kindergarten through fifth grade) from an array of 13 common standardized measures. Outcome data gathered to date in the ongoing, longitudinal study are (a) standardized achievement test data from the Metropolitan Readiness Tests (Version 5, Level 2) administered in April of the kindergarten year (for 2,827 Cohort 1 pupils and 2,889 Cohort 2 pupils) and the Metropolitan Achievement Tests (MAT-6) administered in April of the first grade year (1988) for 1,703 Cohort 1 pupils (those in kindergarten during 1986–1987); and (b) standardized behavior data from teacher ratings on the Hahnemann Elementary School Behavior Rating Scale (Spivack & Swift, 1975) administered in winter 1987, to a random sample of 527 Cohort 1 kindergartners and in April 1988, to 2,570 Cohort 2 kindergartners.

The Hahnemann Scale is a 60-item, standardized rating of children's classroom behavior along 14 dimensions. The positive ones are originality, independent learning, involvement in classroom activities, approach to

teac and productivity with peers. The negative ones are irrelevant talk, socia (over) involvement, negative feelings, intellectual dependency, holding back-withdrawn, critical-competitive, blaming, unreflectiveness, and failure/anxiety.

As the titles of the dimensions suggest, several reflect positive aspects of children's behaviors while others reflect negative aspects of children's behavior. A high score on the positive dimensions reflects positive performance while a low score reflects more negative behavior. A high score on the negative dimensions reflects negative performance while a low score reflects more positive behavior. The number of items related to each dimension varies from 3 to 5; 35 items are rated on a 5-point scale and 25 are rated on a 7-point scale. The number of possible points for each dimension varies from a low of 4 to a high of 35.

In the ongoing longitudinal study, prior preschool attendance and survey data describing the characteristics of many of the preschools attended by the children were reported by the parents. In addition, trained observers provided descriptions of teacher and child behaviors characteristic of each of the kindergartens in the study.

RESULTS

Cognitive Outcomes

We begin with a summary of the cognitive and school-related outcomes as reported in Sheehan et al. (1991).

Full-day versus half-day versus alternate-day kindergarten. Both the retrospective and the longitudinal studies provide clear evidence positively linking participation in full-day kindergarten to test performance, at least through the first grade. The difference between full-day kindergarten and half- or alternate-day kindergarten was evident in standardized test performance (approximately 5 to 10 percentile-point differences favoring full day), fewer grade retentions (17%–55% fewer retentions favoring full day), and lower incidence of Chapter 1 placements (50%–90% fewer Chapter 1 placements favoring full day). Participation in full-day kindergarten was not related to provision of special educational services.

Prior attendance at preschool or day care. Children who attended an early childhood program (preschool or day care) during the year before kindergarten scored approximately 10 percentile points higher on standardized achievement tests such as the Metropolitan Achievement Test (MAT-6) than d children who had not had such experiences. This relationship between school attendance and test performance was evident even at the end of the second grade. The absence of preschool is *not* related to comparatively v performance, as the average pupil without preschool performed at or

somewhat above the national norm on these standardized tests. her, the presence of preschool attendance is related to better-than-average performance on national norms. Children attending preschool were approximately half as likely to be retained (5%) in the early grades when compared to children with no preschool experience and more than half as likely (a 3.5% to 11% range) to receive Chapter 1 services than children with no preschool. There was no relationship between preschool and provision of special educational services.

Gender of the child. On standardized tests through the third grade, girls outperformed boys by 5 to 8 percentile points. Twenty-two percent to 24% fewer girls than boys experienced Chapter 1 placements and 33% to 38% fewer girls than boys were retained in grade. No gender differences were found in the provision of special educational services.

Age at entrance to kindergarten. Summer children who attended kindergarten among the youngest in their class scored lower on standardized tests (5 to 19 percentile points) through at least the first grade, and there are some indications that these differences were apparent even in the third grade. In addition, 14% to 37% of summer children subsequently received Chapter 1 services.

Summer children were most likely to be retained at least once during their elementary years (20% to 25% were retained). In contrast, summer children who delayed kindergarten entry a year after they were actually eligible were least likely to be retained in grade (10% were retained).

Children older than 72 months on October 1 of their kindergarten year were most likely to be in special education placements. This finding is most likely an indication that the possibility of special education placement may have been suspected or known by the parents of these children and may have played a factor in the decision made for their children to enter kindergarten more than 16 months later than their same-age peers.

Behavior Outcomes

Full-day versus half-day versus alternate-day kindergarten. In both cohorts of the prospective study, a clear relationship between kindergarten schedule and children's classroom behavior emerges. Analysis of variance for each cohort shows significant differences related to schedule for originality, independent learning, involvement in classroom activities, productivity with peers, intellectual dependency, failure/anxiety, unreflectiveness, holding back or withdrawn, blaming (Cohort 2 only), and approach to teacher. Each of these differences favored the full-day kindergartners (see Table 2, p. 194). Moreover, with the exception of blaming (Cohort 1 only), no dimensions of children's behavior were more positive for the half-day or alternate-day kindergartners.

Table 'ndergarten Pupils' Reported Behaviors by Kindergarten Schedule

Teacher Perception of Children's Behavior	Schedule			Schedule		
Cohort	Half 1	Alt. 1	Full 1	Half 2	Alt. 2	Full 2
Range of *N*	(262–267)	(195–203)	(53–54)	(996–1001)	(621–630)	(506–532)
† Originality range 4–20	11.1	11.0	12.4*	10.8	10.0	11.6*
† Independent Learning range 5–33	17.0	18.0	19.4*	17.5	17.5	18.5*
† Involvement range 5–27	17.3	18.2	19.0*	17.0	17.5	18.5*
† Productivity w/Peers range 3–21	13.5	14.3	14.2*	13.5	13.5	14.1*
Intellectual Dependency range 4–24	12.6	12.4	10.6*	11.2	11.2	10.4*
Failure/Anxiety range 5–29	12.8	11.2	10.9*	11.0	10.5	10.0*
Unreflectiveness range 3–17	8.0	7.6	6.4*	7.2	6.9	6.6*
Irrelevant Talk range 4–20	8.9	8.5	6.6	8.0	8.0	7.6
Social(Over)Involvement range 4–22	11.2	10.6	9.2	10.0	10.0	9.6
Negative Feelings range 5–27	8.5	7.2	7.2	7.5	7.5	7.5
Holding Back–Withdrawn range 5–35	12.9	12.5	11.8*	11.5	11.5	10.5*
Critical–Competitive range 4–22	8.7	8.0	8.0	8.0	8.0	8.0
Blaming range 4–24	8.2	6.6	6.9	7.2	6.8	6.8*
† Approach to Teacher range 4–24	16.1	15.9	16.3*	15.6	15.2	17.2*

† A high score for each of these items indicates positive behavior.
* Statistically significant ANOVA ($p < .05$) related to schedule of effects.

Prior attendance at preschool or day care. As shown in Table 3, with very few exceptions, kindergartners with prior preschool experience are rated more positively by their teachers than are pupils with no such experience. Statistically significant differences (ANOVAs) are in the areas of originality (Cohort 2 only), independent learning, involvment in classroom activities (Cohort 2 only), intellectual dependency, failure/anxiety (Cohort 2 only), irrelevant talk (Cohort 1 only) and holding back or withdrawn (Cohort 2 only). We should also note that in Cohort 2 kindergarten pupils v preschool experience are also rated by their teachers to be more tive toward teachers and the learning setting, and more critical or

Table 3. Kindergarten Pupils' Reported Behaviors

Teacher Perception of Children's Behavior	Preschool Attendance			
Cohort	No 1	Yes 1	No 2	Yes 2
Range of *N*	(157–160)	(159–164)	(783–833)	(1071–1138)
† Originality range 4–20	11.2	11.1	10.7	11.5*
† Independent Learning range 5–33	17.2	18.6*	17.8	18.4*
† Involvement range 5–27	17.7	18.2	17.3	18.0*
† Productivity w/Peers range 3–21	13.7	14.1	13.8	14.0
Intellectual Dependency range 4–24	12.4	11.9*	11.3	10.7*
Failure/Anxiety range 5–29	12.0	12.1	10.9	10.5*
Unreflectiveness range 3–17	7.7	7.3	7.0	6.9
Irrelevant Talk range 4–20	8.8	8.0*	7.8	7.9
Social(Over)Involvement range 4–22	11.2	10.4	9.9	10.2
Negative Feelings range 5–27	7.9	7.7	7.3	7.6*
Holding Back–Withdrawn range 5–35	12.9	12.0	11.4	10.8*
Critical–Competitive range 4–22	8.7	8.3	7.7	8.1*
Blaming range 4–24	7.5	7.1	6.7	6.9
† Approach to Teacher range 4–24	16.0	5.9	16.1	16.1

† A high score for each of these items indicates positive behavior.
* Statistically significant ANOVA ($p < .05$) related to preschool.

competitive with their peers than are kindergarten pupils with no such preschool experience.

Gender of the child. The impact of gender on kindergartners' classroom behavior as rated by teachers is perhaps the strongest of any variable in this study (see Table 4, p. 196). Kindergarten teachers rated the behavior of girls in both cohorts significantly more positively (using ANOVA testing) than they rated the behavior of boys on *nearly every* behavioral dimension with the exception of originality (a dimension yielding no significant differences for both cohorts) and involvement (Cohort 1 not significant).

1-53

Table 4. Kindergarten Pupils' Reported Behaviors

Teacher Perception of Children's Behavior	Gender			
Cohort	Boy 1	Girl 1	Boy 2	Girl 2
Range of *N*	(266–271)	(245–251)	(1133–1196)	(1155–1180)
† Originality range 4–20	11.1	11.3	11.0	11.1
† Independent Learning range 5–33	17.1	18.2*	17.5	18.3*
† Involvement range 5–27	17.5	18.9	17.3	17.8*
† Productivity w/Peers range 3–21	13.6	14.2*	13.3	14.0*
Intellectual Dependency range 4–24	12.7	12.0*	11.4	10.9*
Failure/Anxiety range 5–29	12.4	11.6*	11.0	10.4*
Unreflectiveness range 3–17	8.2	7.2*	7.5	6.6*
Irrelevant Talk range 4–20	8.8	8.2*	8.5	7.4*
Social(Over)Involvement range 4–22	11.6	10.0*	11.1	9.2*
Negative Feelings range 5–27	8.3	7.4*	8.1	6.9*
Holding Back–Withdrawn range 5–35	13.4	11.9*	11.6	10.8
Critical–Competitive range 4–22	8.5	8.3	8.5	7.6*
Blaming range 4–24	7.9	7.0*	7.4	6.4*
† Approach to Teacher range 4–24	14.9	17.1*	15.1	16.8*

† A high score for each of these items indicates positive behavior.
* Statistically significant ANOVA ($p < .05$) related to gender.

Age at entrance to kindergarten. Statistical testing (ANOVA) confirmed that summer children who attend kindergarten as the youngest of their peer group tend to be viewed by kindergarten teachers in the spring of their kindergarten year as being significantly less original, less independent in learning, less involved, less productive with peers, more intellectually dependent, more prone to anxiety of failure, more unreflective, more prone to irrelevant talk, more holding back and withdrawn, more blaming, and less willing to approach teachers than their older peers. As Table 5 indicates, the children in age groups 3 and 4 (turned 5 from February 1–September 30 of the year before the year they entered kindergarten) were judged by kindergarten teachers to exhibit the more positive behaviors. This finding is consistent with the belief that summer children who enter kindergarten as the oldest of their peer group (those in group 4) perform more positively in kindergarten classes than do the summer children who attend kindergarten in the fall immediately following their fifth birthday. We also recall our earlier discussion of Table 1 indicating that many parents voluntarily delay their summer children's entrance to kindergarten by a full year. As a result, kindergarten

Table 5. Kindergarten Pupils' Reported Behaviors by Age Group—C

Teacher Perception of Children's Behavior	Youngest – Oldest				
	Boy 1[a]	Girl 2	Boy 3	Girl 4[b]	5
Range of *N*	(494–538)	(654–691)	(590–627)	(199–203)	(53–54)
† Originality range 4–20	10.5*	10.9	11.7	11.6	10.5
† Independent Learning range 5–33	17.1*	17.6	18.9	18.9	16.3
† Involvement range 5–27	7.0*	17.4	18.1	18.5	17.2
† Productivity w/Peers range 3–21	3.2*	13.6	14.1	13.9	12.9
Intellectual Dependency range 4–24	11.8*	11.4	10.5	10.7	11.8
Failure/Anxiety range 5–29	11.8*	10.9	10.3	10.2	11.6
Unreflectiveness range 3–17	7.6*	7.2	6.6	6.9	7.4
Irrelevant Talk range 4–20	8.3*	8.2	7.6	8.0	8.6
Social(Over)Involvement range 4–22	10.5	10.2	9.8	10.3	10.7
Negative Feelings range 5–27	7.6	7.5	7.4	7.9	8.6
Holding Back–Withdrawn range 5–35	12.3*	11.7	10.1	10.4	12.8
Critical–Competitive range 4–22	8.1	8.0	8.1	8.5	8.1
Blaming range 4–24	7.2*	6.9	6.6	6.9	8.2
† Approach to Teacher range 4–24	15.9*	15.9	16.3	16.5	16.4

† A high score for each of these items indicates positive behavior.
* Statistically significant ANOVA ($p < .05$) related to entrance age.
[a] Summer children attending kindergarten.
[b] Summer children held back.

teac[illegible]ave 4 to 6 "young" summer children out of 25 children rather than the approximately 8 such children who would be enrolled if all children attended kindergarten when they were age-eligible. The smaller number of young summer children in a class makes the immaturity of the young summer children all the more apparent. Findings for Cohort 1 children are similar to those reported above for children in Cohort 2 (Table 5). Tabular data is omitted in the interest of brevity.

DISCUSSION

Our conclusions about the impact of full-day kindergarten are subject to a few qualifications. To address the issue of the comparability of the "educational program" experienced by children in each of the kindergarten classes, we verified that each class was receiving the same quality program by conducting a review of the courses of study on file, lesson plans, and posted schedules. Further, we carefully observed each classroom three times during the kindergarten year focusing upon specific teacher and child behaviors and the quality of the learning environment. Analysis of these data is not complete. Preliminary data suggest that children experienced essentially the same type of kindergarten experience across districts and by schedule. Two exceptions are noted. Half-day children spent a greater percentage of their time in teacher-led large groups, and children in both full- and alternate-day kindergartens spent a greater percentage of their time in active free play. These findings are of interest because early childhood professionals have expressed concern that districts that adopt full-day kindergarten may expose children to more academic programs with less time devoted to play (Olsen & Zigler, 1989).

Only a small percentage of the pupils in the study were charged tuition for full-day kindergarten. In most cases, the provision of full-day kindergarten was at total cost to the school district, and enrollment in full-day kindergarten was made on a space-available basis to any parent requesting such a schedule.

A second qualification pertains to potential economic discrepancies between the children attending the different kindergarten schedules. If families had to pay considerable tuition for children to attend full-day kindergarten the data could be biased in favor of economic circumstances.

Our findings in favor of full-day versus half-day and alternate-day pupils, are probably an underestimate of the true impact of full-day kindergarten. Previous research (Sheehan, 1988) has indicated that 56% of the pupils in half-day kindergarten in Ohio spend at least some portion of the *remainder* of their day in child care programs outside the home. At least some of these child care programs are likely to have an educational or behavioral component serving to complement the educational and behavioral impact of half-day kindergarten in a fashion similar to the full-day programs.

The consistent findings for the full-day versus half-day versus alternate-day kindergarten schedule, across two cohorts of children and across the many dimensions of the Hahnemann Elementary School Behavior Rating Scale, provide strong support for the effectiveness of the full-day kindergarten on children's classroom behavior. Children who have attended full-day kindergarten are rated more positively by their teachers than are children in half- or alternate-day programs. However, we also note that the averages for all three groups reported in Table 2 are well within the normal range of behaviors expected during the school year (Spivack & Swift, 1975), but that full-day pupils are rated as having exhibited more positive behavior than did the pupils in half-day or alternate-day kindergarten.

Any conclusions about the relationship between school performance and prior preschool attendance are threatened by the socioeconimic correlates of preschool attendance. In most cases parents must pay for early childhood experiences (Head Start is a notable exception to this statement), and the ability to pay for preschool (or day care) may be a reflection of a larger ability to provide home environments that are rich in educational stimuli. Our results of preschool effects on behavior ratings share this limitation. Although our selection of school districts represented a balance of socioeconomic circumstances, we were not able to gather data on the socioeconomic circumstances of our subject population. However, subset analyses of school districts that represent fairly low socioeconomic circumstances and fairly urban socioeconomic circumstances evidence an even stronger, positive relationship between school performance and prior preschool attendance.

Just as we were unable to control for the socioeconomic circumstances of the subject population, so too were we unable to control for or assess the quality of the preschool experience. It may well be that a "good" preschool experience is a powerful predictor of future elementary school success for *all* children.

The fact that more boys than girls are diagnosed as special education pupils (Miles, 1986) is not surprising to special educators, and a variety of explanations has been offered for that pattern (Eme, 1979; Oakland & Laurent, 1987). The finding mentioned above that boys are more likely to be in Chapter 1 programs and are more likely to be retained *is* surprising, troubling, and perhaps not well known to parents and public school personnel.

The classroom behavior results obtained from both cohorts of the prospective study are similar to that reported by Spivak and Swift (1975) in the manual for the Hahnemann Elementary School Behavior Rating Scale. This scale is not designed to be free of gender bias; rather, it reflects the "oft-reported indications that boys are more prone than girls to exhibit behavioral difficulties" (unpaged, Spivack & Swift, 1975). Numerous qualifications of the impact of kindergarten entrance age must be noted. First, we do not know the reasons that certain summer children's entrance to kindergarten was delayed and others went on to attend, reasons that may be related to

chil. 's subsequent school performance. Second, we do not know what types of educational experiences may have been provided to children during he year in which they did not attend public school kindergarten (were held back). Did they experience an additional year of preschool (as 4-year-olds and 5-year-olds)? Were they in private kindergarten? Such experiences may have influenced their subsequent school performance. Third, there is no information available on the economic factors facing families as they make a decision to enroll or not enroll a summer child in kindergarten. For many families, enrollment in kindergarten may be based upon the need for a parent to resume work, either full- or part-time.

Age alone is not a good rationale for delaying children's entrance to kindergarten. Elementary school classes will always encompass at least a 12-month span of children, some older, some younger. Rather, we interpret our data to indicate at least one group of children at greater risk for school failure than others. Educational intervention (such as Chapter 1—see Slavin, 1987) and increased sensitivity of families and educators to the needs of these young children might reduce that risk status.

CONCLUSIONS

Policymakers with concern for early education are faced with putting together a puzzle that is missing a number of pieces. First, a quality preschool education has long-lasting beneficial outcomes for disadvantaged children. But, does a quality preschool experience benefit *all* children? If so, are the benefits significant enough to offset the enormous expense? Second, recent practice in many states is for children to attend half-day kindergarten programs. But, social and economic changes sending more than 50% of mothers with children into the workforce dictate a growing need for full-day kindergarten. Does full-day kindergarten benefit the children as much as it benefits their working parents? Are full-day kindergarten programs likely to be developmentally appropriate, with play as the basis for learning, or will they be extended versions of the academic pressure-cooker approach that has resulted from the downward extension of the first-grade curriculum? Early childhood professionals resist promoting full-day kindergarten fearing the latter. Third, when should children start school? If preschool lives up to its billing and states are able financially to afford public preschool, then children will be able to start school at age 3. Transition to kindergarten could be made based on an accumulation of developmental data for the individual child rather than a child's age.

We feel that this study has provided some of the missing puzzle pieces. were surprised at the almost total absence of statistical interactions in our data. Each of the factors we presented operates by itself as a powerful ain effect to fill in a piece of the puzzle. First, preschool seems to have a

balanced beneficial effect (academic and behavioral) for all . Jren in our study. Legislators and state departments of education may now be encouraged to provide seed money to school districts to begin the process of providing quality preschools for all children, particularly those of working parents. With the availability of federal funds for preschool handicapped children as the incentive, school districts could consider developing integrated programs for handicapped and nonhandicapped children, with parents of nonhandicapped children paying modest tuition to support the extended funding needs.

Second, full-day kindergarten was of benefit to children in the study. The benefits seem to last well into the second grade. At the outset of the study we were cautioned by many teachers to consider outcomes other than academics. Many felt that sole reliance upon standardized testing for kindergarten-aged children was inappropriate. Others worried that kindergartens in the study would not be comparable—that the deck was being stacked against the half-day kindergarten where teachers had fewer hours to teach academic skills on which the children would be tested. Our observation data indicate that the different kindergartens were in fact quite comparable; teachers and children in half-day kindergartens spent proportionately the same time involved in the same types of activities as did teachers and children in full-day programs. Contrary to expectation, however, the teachers in full-day kindergartens tended to allow children proportionately more time for active free play. Looking beyond the academic outcomes, we found that teachers have very different feelings about kindergartners in full-day programs. Although children in the study were typically perceived by teachers as having behavior well within the normal range, regardless of schedule type, those in full-day kindergarten were seen by their teachers to be significantly better behaved on dimensions valued by educators. Specifically, full-day children were more involved, showed more originality and independent learning, and were less likely to be dependent, shy, and withdrawn than their half-day and alternate-day counterparts. These very positive behavioral outcomes coupled with the significant achievement gains and subsequent school outcomes may help allay the fears of early childhood professionals and might encourage more in-depth experiments with developmentally appropriate full-day kindergarten, while policymakers support cost effectiveness studies to verify the real benefits of full-day kindergarten.

Third, it is quite risky to be a "Summer Child." Attending kindergarten as one of the youngest children in the class is directly related to increased school failure, leading to increased possibilities for retention and referrals for special educational services (i.e., Chapter 1). These outcomes are extremely expensive for school districts. Reducing the numbers of children at risk for these problems is not accomplished by changing the school-age entrance date. Such moves only solve the problem for one year. There will

always 'young" group. The solution lies elsewhere, possibly in flexible entrance dates linked to preschool attendance and well-designed developmental screening and assessment programs (Meisels, 1987). School personnel need to have the capability to identify individual differences in young children and link the assessment of those differences with a developmentally appropriate curriculum (Bagnato, Neisworth, & Munson, 1989). All children's potential to succeed would be enhanced.

REFERENCES

Bagnato, S.J., Neisworth, J.T., & Munson, S.M. (1989). *Linking developmental assessment and early intervention: Curriculum-based prescriptions* (2nd ed.). Rockville, MD: Aspen.

Barnett, W.S., & Escobar, C.M. (1987). The economics of early intervention. *Review of Educational Research, 4,* 387–414.

Campbell, S.M. (1985, March). *Kindergarten entry age as a factor in academic failure.* Paper presented at the Annual Convention of the American Association of School Administrators, Dallas, TX. (ERIC Document Reproduction Service No. ED 256 495)

Coates, S. (1974). Sex differences in field independence among preschool children. In R.C. Friedman, R.M. Richart, & R.L. Vande Wiele (Eds.), *Sex differences in behavior* (pp. 259–274). New York: Wiley.

Copple, C.E., Cline, M.G., & Smith, A.N. (1987). *Path to the future: Long term effects of Head Start in the Philadelphia school district.* Washington, DC: U.S. Department of Health and Human Services.

Cryan, J.R. (1986). *Kindergarten scheduling alternatives: A review of research.* Columbus: Ohio Department of Education, Division of Early Childhood.

Darlington, R.B., Royce, J.M., Snipper, A.S., Murray, H.W., & Lazar, I. (1980). Preschool programs and later competence of children from low-income families. *Science, 208,* 202–204.

Diamond, G.H. (1983). The birthdate effect: A maturational effect. *Journal of Learning Disabilities, 16,* 161–164.

Dwyer, C.A. (1973). Sex differences in reading: An evaluation and a critique of current methods. *Review of Educational Research, 43,* 455–461.

Elkind, D. (1988). *The hurried child: Growing up too fast too soon* (2nd ed.). Reading, MA: Addison-Wesley.

Eme, R. (1979). Sex differences in childhood psychopathology: A review. *Psychological Bulletin, 86,* 574–593.

Gullo, D.F. (1990). The changing family context: Implications for the development of all-day kindergartens. *Young Children, 45*(4), 35–39.

Gullo, D.F., Bersani, C., Clements, D., & Bayless, K. (1986). A comparative study of "all-day," "alternate-day," and "half-day" kindergarten schedules: Effects of achievement and classroom social behaviors. *Journal of Research in Childhood Education, 1,* 87–94.

Harris, S., & Harris, L.B. (1986). *The teacher's almanac.* New York: Facts on File Publication.

Helmich, E., & Wasem, L. (1985, January). *Class sizes for kindergarten and primary grades: A review of the research* (Research Report). Springfield: Illinois State Board of Education. (ERIC Document Reproduction Service No. ED 260 827)

Karweit, N.L. (1987). *Full or half day kindergarten—Does it matter?* (Report No. 11). Baltimore, MD: The Johns Hopkins University Center for Research on Elementary and Middle Schools.

Karweit, N.L. (1988). Quality and quantity of learning time in preprimary programs. *Elementary School Journal, 89,* 119–133.

Karweit, N.L. (1989a). Effective kindergarten programs and practices for students at risk of academic failure. In R.E. Slavin, N.L. Karweit, & N.A. Madden (Eds.), *Effective programs for students at risk* (pp. 103–142). Boston, MA: Allyn & Bacon.

Karweit, N.L. (1989b). Preschool programs for students at risk for school failure. In R.E. Slavin, N.L. Karweit, & N.A. Madden (Eds.), *Effective programs for students at risk* (pp. 75–102). Boston, MA: Allyn & Bacon.

Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Meisels, S.J. (1987). *Developmental screening in early childhood: A guide* (rev. ed.). Washington, DC: National Association for the Education of Young Children.

Miles, D.L. (1986). Why do more boys than girls receive special education? *Contemporary Education, 57*(2), 104–106.

Nurss, J.R., & Hodges, W.L. (1982). Early childhood education. In H.E. Mitzell (Ed.), *Encyclopedia of Educational Research* (vol. 1, pp. 489–513). New York: Macmillan.

Oakland, T., & Laurent, J. (1987). Sex differences in learning abilties. In C.R. Reynolds & L. Mann (Eds.), *Encyclopedia of Special Education: A reference for the education of the handicapped and other exceptional children and adults.* (Vol. 3, pp. 1423–1424). New York: Wiley.

Olsen, D., & Zigler, E. (1989). An assessment of the all-day kindergarten movement. *Early Childhood Research Quarterly, 4,* 167–186.

Plomin, R., & Foch, T.T. (1981). Sex differences and individual differences. *Child Development, 52,* 386–388.

Randolph, K.M. (1986). A study of the fourth and sixth grade achievement of young and old school entrants with different kinds of preschool experience. (Doctoral dissertation, University of Connecticut, 1986). *Dissertation Abstracts International, 47,* 3271A.

Schweinhart, L.J., Weikart, D.P., & Larner, M.B. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly, 1,* 15–45.

Sheehan, R. (1988). *1986–1987 Annual Report: Ohio Department of Education: Preschool kindergarten longitudinal research study.* Columbus: Ohio Department of Education, Division of Early Childhood.

Sheehan, R., Cryan, J.R., Wiechel, J., & Bandy-Hedden, I. (1991). Factors contributing to success in elementary schools: Research findings for early childhood educators. *Journal of Research in Childhood Education, 6,* 1–10.

Shepard, L.A., & Smith, M.L. (1986). Synthesis of research on school readiness and kindergarten retention. *Educational Leadership, 44,* 78–86.

Slavin, R.E. (1987). Making chapter 1 make a difference. *Phi Delta Kappan, 69*(2), 110–119.

Spivack, G., & Swift, M. (1975). *Hahnemann Elementary School Behavior Rating Scale.* Philadelphia, PA: Hahnemann University.

Uphoff, J.K., & Gilmore, J. (1985). Pupil age at school entrance: How many are ready for success? *Educational Leadership, 43,* 86–90.

Uphoff, J.K., & Gilmore, J. (1986). Viewpoint 2: Pupil age at school entrance—How many are ready for success? *Young Children, 41*(2), 11–16.

Walsh, D.J. (1989). Changes in kindergarten: Why here? Why now? *Early Childhood Research Quarterly, 4,* 377–391.

TEACHER PERFORMANCE

Accountability in the Classroom

Research Summary

Study after study shows the single most important element in a child's education is the quality of his or her teacher. A teacher's ability to instruct well is more important than class size and more important than the level of other students in the classroom. The results of a good teacher in one grade can last at least two grades later, while with one bad teacher a student can lose a full level of achievement in a single school year.

What makes a teacher effective and how can we ensure teachers are accountable for the results they produce in the classroom? There is no single answer, but the research points to strong verbal and math skills for teachers in all grades, deep knowledge in teachers' content area, and technological skills. Though Kansas ranks in the top 15 states for teachers with majors in content areas, many studies point out majors do not necessarily require in-depth knowledge of subject matter. Instead, they often rely on survey courses.

Glasscock-Tanner "Years of Promise -- Kansas' Keys to Success" Education Plan

While researchers are still evaluating various approaches to improving teacher performance, there are steps Kansas should take now to make schools, teachers and teacher education programs accountable for their results.

Accountability to Parents:

- Every teacher must have a Kansas Bureau of Investigation background check in order to be licensed to teach in Kansas.
- Every student's parents deserve to receive the Kansas Department of Education school report card which details the strengths and weaknesses of their child's school.
- Every student's parents deserve to know how their child's teacher is licensed.
- Students will not have first-year teachers without adequate professional support. Every first-time teacher will have an on-site mentor teacher during their first year teaching. The KSDE will set standards for and provide mentor training in order to begin the program in FY2001. Mentors will receive $1,000 for a school year of service.
- Students will benefit from teachers who have their quality of teaching consistently reviewed by their peers.

- Students will not have as many out-of-discipline teachers. The state will provide 100 scholarships of $5,000 to students preparing to teach in high-need content areas or geographic areas in the state.

- Teachers will have incentives to improve their teaching. National Board Certified teachers – the most rigorous certification program in the nation – will receive a total of $10,000 for improving the quality of their teaching. They will receive $1,000 yearly during the 10-year license being implemented by the KSDE.

Higher Education Accountability:

- The state's higher education system will prepare the state's 7th - 12th grade teachers with strong content backgrounds. By 2003, each college graduate from an Education program will pass subject matter tests in their areas of expertise in order to be licensed. The higher education institutions will administer these tests and any out of state teacher to be licensed must take the content test from any of the approved higher education institutions in Kansas.

- Kansas education schools must stand behind the quality of the teachers they prepare. Each higher education institution must guarantee their teachers are ready to teach in a Kansas classroom. If, upon a school's determination, a teacher is not performing, the higher education institution must provide necessary supplemental training for a graduate prior to the teacher obtaining a five-year license.

- Higher education teaching schools must report to the public their rate of placing their graduates in teaching positions. In addition, they must report to the public the percentage of their graduates who pass the licensing test.

Technology in the Classroom:

- To prepare teachers to bring technology into the classroom, the Board of Regents will require departments and colleges of education to integrate technology in the Education curriculum. The Kansas Board of Education will include this integration of technology in its criteria for accrediting teacher training programs.

TEACHER PERFORMANCE RESEARCH HIGHLIGHTS

"'The difference between a good and a bad teacher can be a full level of achievement in a single school year'."

"Accountability measures for colleges and universities that prepare teachers.... [mean that states] need to decide on what intending teachers need to know in their subjects and hold academic departments accountable for getting them there before they graduate."

"Parents deserve to know when their children are being taught science by history majors or history by physical education grads.... but nowhere has there been a systematic way of letting all parents know that their child's teacher has enough background in the subject to teach it so their students will understand it."

"We can produce the highly qualified teachers that we need by combining high entry standards [and] rich incentives like generous scholarships..."

Kati Haycock
"Good Teaching Matters ... A Lot"
Thinking K-16

"The results show that teacher effects are dominant factors affecting student academic gain and that the classroom context variables of heterogeneity among students and class sizes have relatively little influence on academic gain. Thus, a major conclusion is that teachers make a difference."

S. Paul Wright, Sandra P. Horn and William L. Sanders
"Teacher and Classroom context Effects on Student Achievement: Implications for Teacher Evaluation"
Journal of Personnel Evaluation in Education

"An expanded pool of high quality teachers is ensured by a comprehensive approach to enlisting the best.... A mandated salaried, mentored induction year gives new teachers classroom responsibility with intensive support. All teachers [should] have the opportunity to become nationally certified."

Lowell Milken, Milken Family Foundation
A Matter of Quality: A Strategy for Assuring the High Caliber of America's Teachers

"Every child should be able to count on having a teacher who has a solid general education, who possesses deep subject area knowledge, and who has no record of misbehavior. The state has an obligation to ensure that all prospective teachers meet this minimal standard. Thus states should perform background checks on candidates for teaching positions."

"The Teachers We Need and How to Get More of Them"
Thomas B. Fordham Foundation

"What little systematic research has been done suggests that classroom technology can raise student achievement and even improve the overall learning environment in schools – but only when it is placed in the right hands and used in the right ways.... Students whose teachers had professional development in computers outperformed – by more than one-third of a grade level – students whose teachers did not."

The Progress of Education Reform 1998
Education Commission of the States

A PUBLICATION OF THE EDUCATION TRUST
Vol. 3, Issue 2

GOOD TEACHING MATTERS ... A LOT

by Kati Haycock
Director, The Education Trust

Parents have always known that it matters a lot which teachers their children get. That is why those with the time and skills to do so work very hard to assure that, by hook or by crook, their children are assigned to the best teachers. (That is also at least part of the reason why the children of less skilled parents are often left with the worst teachers, but more on that later.)

Professional educators typically reject these notions. When parents ask for their children to be assigned to a particular teacher, or to be moved out of the classroom of another, most principals counsel them not to worry. "Your child will learn what he or she needs to from any of our teachers."

Recent research from Tennessee, Texas, Massachusetts and Alabama proves that parents have been right all along. They may not always know which teachers really are the best, but they are absolutely right in believing that their children will learn a lot from some teachers and only a little from others—even though the two teachers may be in adjacent classrooms. "The difference between a good and a bad teacher can be a full level of achievement in a single school year," says Eric Hanushek, the University of Rochester economist notorious for macroanalyses suggesting that virtually nothing seems to make a difference.[1]

TEACHER EFFECTS: TENNESSEE

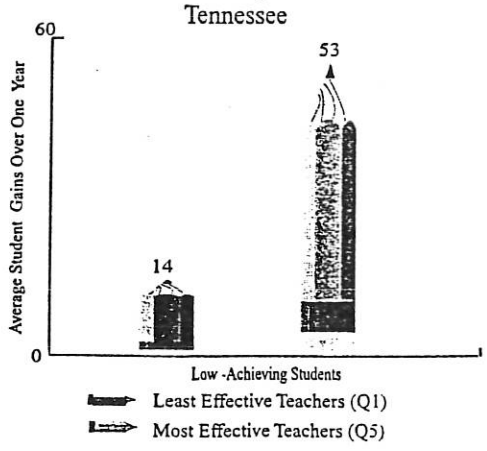
Tennessee is one of the few states with data systems that make it possible to tie teachers to achievement in their classrooms. Moreover, the state's value-added approach for assessing student achievement allows observers to look at the gains students make during a particular school year.

William L. Sanders, director of the Value-Added Research and Assessment Center at the University of Tennessee, Knoxville, has studied these data extensively. By grouping teachers into quintiles based on their effectiveness in producing student learning gains, his work allows us to examine the impact of teacher effectiveness on the learning of different types of students, from low- to high-achievers.

The chart adjacent shows the effect teachers from different quintile levels have on low-achieving students. On average, the least effective teachers (Q1) produce gains of about 14 percentile points during the school year. By contrast, the most effective teachers (Q5) posted gains among low-achieving students that averaged 53 percentile points.

The Tennessee data show dramatic differences for middle- and high-achieving groups of students, too.

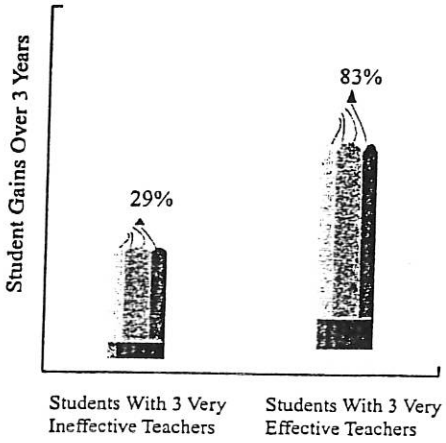
The Effect of Different Teachers On Low-Achieving Students
Tennessee



Sanders, William L. and Rivers, Joan C. "Cumulative And Residual Effects of Teachers on Future Student Academic Achievement," 1996, Table 1, p. 9.

1-61

Cumulative Effects of Teacher Sequence on Fifth Grade Math Scores: Tennessee



Sanders, William L. and Rivers, Joan C., "Cumulative And Residual Effects of Teachers on Future Student Academic Achievement," 1996, Figure 1, p.12

For example, high-achieving students gain an average of only 2 points under the direction of Q1 (least effective) teachers but an average of 25 points under the guidance of Q5 (most effective) teachers. Middle achievers gain a mere 10 points with Q1 teachers but in the mid-30s with Q5 teachers.

There is also considerable evidence that, at least in Tennessee, the effects of teachers are long-lived, whether they advance student achievement or squash it. Indeed, even two years after the fact, the performance of fifth-grade students is still affected by the quality of their third-grade teacher. The chart above shows the examples of different patterns of teacher effectiveness for one metropolitan system.

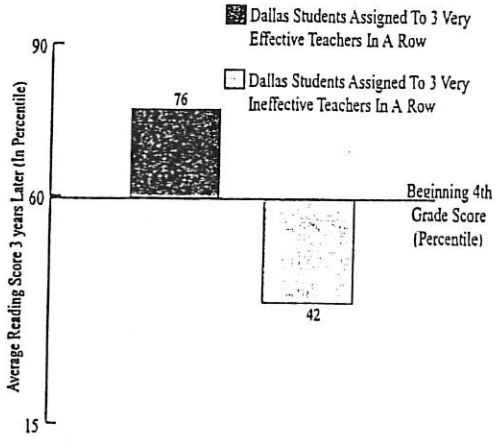
As Sanders points out, students whose initial achievement levels are comparable have "vastly different academic outcomes as a result of the sequence of teachers to which they are assigned."[2] Differences of this magnitude—50 percentile points—are stunning. As all of us know only too well, they can represent the difference between a "remedial" label and placement in the "accelerated" or even "gifted" track. And the difference between entry into a selective college and a lifetime at McDonald's.

TEACHER EFFECTS: DALLAS

A variety of recent studies in Texas show similar differences in achievement between students taught by teachers of differing quality. Borrowing from some of Sanders's techniques, researchers in the Dallas Independent School District recently completed their first-ever study of teacher effects on the ability of students to perform on assessments. In sharing their findings, Robert Mendro, the district's executive director of institutional research, said, "what surprised us the most was the size of the effect."[3]

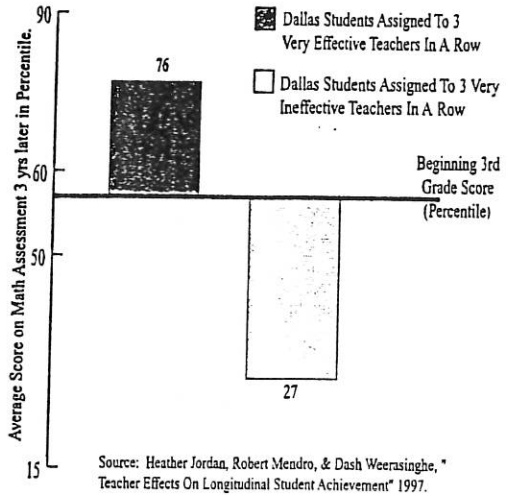
For example, the average reading scores of a group of Dallas fourth graders who were assigned to three highly effective teachers in a row rose from the 59th percentile in fourth grade to the 76th percentile by the conclusion of sixth grade. A fairly similar (but slightly higher achieving) group of students was assigned thr consecutive ineffective teachers and fell from the 60th percentile in fourth grade to the 42nd percentile by the end of sixth grade. A gap of this magnitude—more than 35 percentile points—for students who started off roughly the same is hugely significant.

Effects On Students' Reading Scores In Dallas (Grades 4-6)



Source: Heather Jordan, Robert Mendro, & Dash Weerasinghe, "Teacher Effects On Longitudinal Student Achievement" 1997.

Effects On Students' Math Scores In Dallas (Grades 3-5)



Source: Heather Jordan, Robert Mendro, & Dash Weerasinghe, "Teacher Effects On Longitudinal Student Achievement" 1997.

The impact of teacher effectiveness is also clear in mathematics. For example, a group of beginning third-graders in Dallas who averaged around the 55th percentile in mathematics scored around the 76th percentile at the end of fifth grade after being assigned to three highly effective teachers in a row. By contrast, a slightly higher achieving group of third graders—averaging around the 57th percentile—were consecutively taught by three of the least effective teachers. By the conclusion of fifth grade, the second group's percentile ranking had fallen to 27th. This time the youngsters, who had scored nearly the same as beginning third-graders, were separated by a full 50 percentile points just three years later.

TEACHER EFFECTS: BOSTON

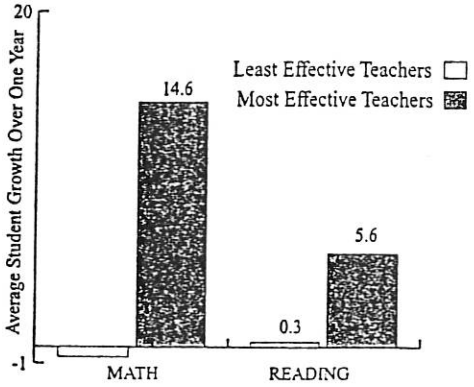
The Boston Public Schools are taking a serious look at factors that influence student learning, including the effectiveness of their teachers. A recently released study by Bain and Company conducted on behalf of the district shows the correlation between high school teachers and their students' academic growth in math and reading. The authors examined classrooms of BPS tenth-graders whose average scores were approximately the same and charted their progress over the year by teacher. The differences were dramatic. In reading, they found that although the gains of students with the top third teachers were slightly below the national median for growth (5.6 on average compared to 8.0), the students with teachers from the bottom third showed virtually no growth (0.3). The math results were even more striking. The top third teachers produced gains on average that exceeded the national median (14.6 to 11.0 nationally), whereas the bottom third again showed virtually no growth (-0.6).

Altogether, this means that one-third of BPS teachers are producing six times the learning seen in the bottom third. As one frustrated headmaster put it, "About one-third of my teachers should not be teaching."

WHAT MAKES FOR TEACHER EFFECTIVENESS?

None of these studies has yet advanced to the obvious next step: identifying the qualities that make for an effective teacher. But other researchers have used Texas's extensive database on both teachers and students to examine the impact of specific teacher characteristics on student achievement. Together with work from Alabama and North Carolina, this research helps us to get underneath the matter of teacher effectiveness.

Boston Students With Effective Teachers Showed Greater Gains



Source: Boston Public Schools, "High School Restructuring," March 9, 1998.

1. Strong Verbal and Math Skills

The first thing that is clear when you look across the various studies is the critical importance of strong verbal and math skills. Harvard's Ronald F. Ferguson, for example, has looked closely at the relationship between student achievement and teacher performance on a basic literacy examination (the Texas Examination of Current Administrators and Teachers, which was administered to all teachers and administrators in Texas in 1986). Ferguson found a significant positive relationship between teacher test scores on TECAT and student scores on the Iowa Test of Basic Skills (ITBS), with higher scoring teachers more likely to produce significant gains in student achievement than their lower scoring counterparts. Indeed, a change of one standard deviation in a district's teacher scores produced a corresponding change of .17 standard deviation in student scores, when other differences were controlled.[4]

Students' test scores can mean the difference between a remedial label and the gifted track — or between entry into a selective college and a lifetime at McDonald's.

Ferguson got similar results in an analysis of the impact of teacher and classroom qualities on student achievement scores in Alabama. As in the Texas studies, he found a strong positive relationship between teacher test scores (in this case, ACT scores) and student achievement results.[5]

2. Deep Content Knowledge

There is also considerable research showing how important teachers' content knowledge is to their effectiveness with students, especially at the middle and senior high school levels. The data are especially clear in mathematics and science where teachers with majors in the fields they teach routinely get higher student performance than teachers who did not. Goldhaber and Brewer examined this relationship using data from the National Educational Longitudinal Study of 1988 (NELS), an ongoing survey of individuals who were in eighth grade in 1988. Goldhaber and Brewer found a significant positive relationship between teachers' degrees and students' achievement in technical subjects. They concluded that "in mathematics and science, it is the teacher subject-specific knowledge that is the important factor in determining tenth-grade achievement."[6]

The data are less clear in English and social studies; in these subjects students taught by majors don't show consistently better scores than students taught by teachers who majored in something else. However, other evidence suggests that content is no less important in these two disciplines. For example, a recent study in Hawaii asked social studies teachers to rate their own level of understanding about various historical periods and teaching methods, then compared teacher expertise to student achievement. Not surprisingly, there was an almost perfect match: students performed best in the domains where teachers indicated the most expertise.[7]

3. Teaching Skill?

All of this seems to beg the question: what about teaching knowledge and skills? Is content knowledge really sufficient for effective teaching? Clearly not. One only has to spend a few semesters in higher education to see that the deep content knowledge inherent in the Ph.D. doesn't necessarily lead to effective teaching.

That said, the large-scale studies we have reviewed are not particularly helpful in identifying ways to quantify teaching expertise. Neither education courses completed, advanced education degrees, scores on professional knowledge sections of licensure exams nor, interestingly, years of experience seem to have a clear relationship to student achievement. Perhaps the work going on at the National Board for Professional Teaching Standards or Lee Shulman's work on "pedagogical content knowledge" at the Carnegie Foundation for the Advancement of Teaching will advance our understanding of—and options for developing and measuring—teaching knowledge and skill.

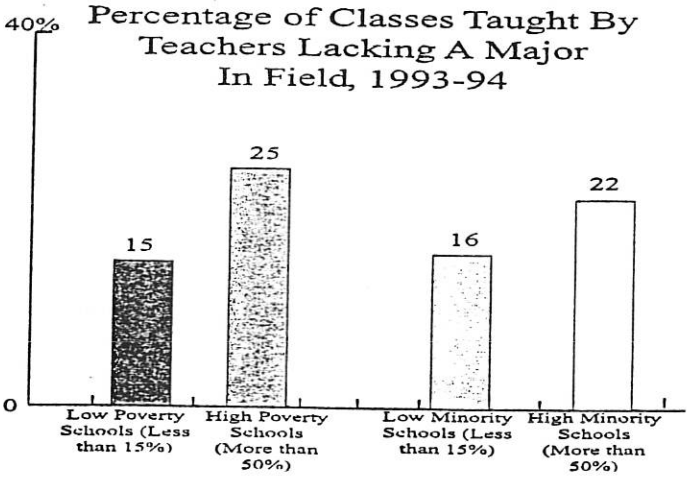
In the meantime, we suggest that educational leaders not get sidetracked: there is more than sufficient evidence about the importance of deep content knowledge and strong verbal skills to serve as a foundation for immediate action. At the very least, we know

1-64

enough to call the question with faculty in the arts and sciences, who, after all, are responsible for developing both content knowledge and verbal skills among intending teachers. It is also enough to justify a second look at hiring and assignment criteria. If good teachers matter, we need to be sure that we are getting the best we can.

INEQUITIES IN DISTRIBUTION

Our emerging understanding of the critical importance of good teachers has especially profound implications for poor and minority youngsters. For no matter how quality is defined, these youngsters come up on the short end. While the teaching force in high-poverty and high-minority communities certainly includes some of the most dedicated and talented teachers in the country, the truth is that these teachers are vastly outnumbered by under- and, indeed, unqualified colleagues.



Source: Richard Ingersoll, University of Georgia, Unpublished, 1998.

These patterns are clear in national data tabulations on out-of-field teaching specially prepared for the Education Trust earlier this year by Richard Ingersoll, a professor at the University of Georgia. As is evident in the table above (as well as in the state tabulations on pp. 8-9) minority and poor youngsters—the very youngsters who are most dependent on their teachers for content knowledge—are systematically taught by teachers with the least content knowledge.

Similar inequities show up at all grade levels in the state-level studies described above, and many more. For example, in Tennessee, black students are almost twice as likely to be taught by ineffective "Q1" teachers as are white children, and are considerably less likely to be taught by the most effective teachers.

The patterns look quite similar in Texas, where, according to researchers John Kain and Kraig Singleton, African American and Latino children are far more likely to be taught by teachers who scored poorly on the TECAT examination. Indeed, as the percentage of non-white children in the school increases, the average teacher score declines.[8] Finding the same patterns in his analysis, Ferguson wrote that "[i]n Texas, and certainly in other places too, attracting and retaining talented people with strong skills to teach in the districts where black students are heavily represented is part of the unfinished business of equalizing educational opportunity."[9]

RACE MORE THAN CLASS?

Contrary to the assumptions that many people may make, inequities in the distribution of teacher expertise are not driven wholly by finances. If they were, we would expect that poor minority children would have teachers of about the same quality as poor white children. But such is not always the case.

In their analysis of Texas data, Kain and Singleton found disturbing differences. Poor white children, it turns out, appear to have a higher likelihood of having well qualified teachers than poor black children.[10]

Similar patterns are evident in teacher quality data from other states. In the chart on pages 8 and 9, for example, it is clear that students who attend predominantly minority secondary schools in Virginia are more likely to be taught by underqualified teachers than stu-

continued page 10

1-65

State Investment in Well-Prepared Teachers

The most important educational investment a state can make is in highly qualified teachers. When teachers have too little knowledge of the subjects that they teach, their students are denied the most basic learning resource. There are several ways to examine teacher quality. This chart shows one: the percentage of secondary school classes taught by teachers who lack a college major in the subject area.

The chart shows, by state:

- the overall percentages of classes taught by teachers who do not have a major in the subject that they teach; and
- the percentages of classes taught by teachers who do not have a major in the subject that they are teaching in high-poverty schools/high-minority schools (schools in which more than 50% of the students are low-income or non-white) vs. low-poverty schools/low-minority schools (schools in which fewer than 15% of the students are low-income or non-white).

In reviewing the chart, the reader will see a stark and troubling pattern: low-income students and students of color are less likely than other students to be taught by teachers with a college major in the subject area that they are teaching.

The data used to build this chart are drawn from the Schools and Staffing Survey conducted by the National Center for Education Statistics (NCES) in school year 1993-94. Richard Ingersoll of the University of Georgia conducted the analysis. While the Schools and Staffing Survey is large scale, in some states the data are inadequate to support stable estimation for certain kinds of schools so we have not printed a percentage. There are other cases where the sample meets normal standards, but the Education Trust staff cautions the reader with an "*" that these samples are "on the smallish side" and advises further research.

We have ranked states on the overall quality of their teachers. The fewer underqualified teachers, the better the rank. We also rank states according to disparity in assignment of underqualified teachers. "Disparity by poverty", for example, is the difference between the percentage of classes in high- and low-poverty schools that are taught by underqualified teachers.

Percentage of secondary school classes taught by teachers lacking a major in field by state, 1993-1994

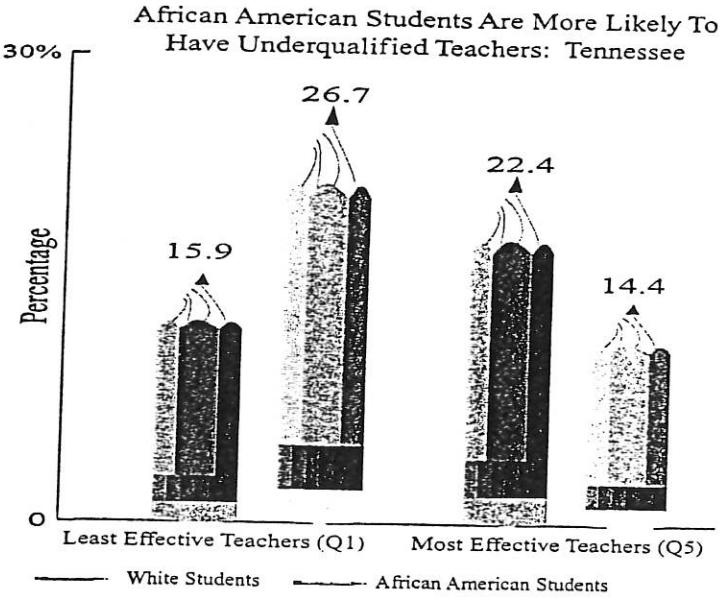
	Overall (Rank)	By School Poverty			By School Minority Population		
		Low	High	(Rank)	Low	High	(Rank)
Alabama	17% (21)	14%	23%	(19)	17%	16%	(7)
Alaska	29 (51)	20	48	(33)	23	41	(29)
Arizona	26 (49)	17	37	(31)	26	29	(13)
Arkansas	14 (8)	11	11	(5)	15	14	(7)
California	27 (50)	28	29	(6)	26 *	27	(12)
Colorado	20 (33)	17			20	24 *	(15)
Connecticut	13 (6)	12			12	15	(13)
Delaware	24 (43)					25	
D.C.	17 (21)					17	
Florida	18 (28)	22	19	(4)	24	17	(2)
Georgia	21 (36)	15	33	(29)	25	19	(3)
Hawaii	24 (43)	13				25	
Idaho	20 (33)	15			19		

1-66

	Overall (Rank)	By School Poverty			By School Minority Population		
		Low	High	(Rank)	Low	High	(Rank)
Illinois	16% (15)	13%	24%	(24)	13%	25%	(25)
Indiana	13 (6)	10	15 *	(11)	12	26 *	(27)
Iowa	12 (4)	7	17	(23)	12		
Kansas	15 (12)	15			14		
Kentucky	24 (43)	20	29	(19)	24		
Louisiana	23 (40)	9	28	(30)	21	21	(10)
Maine	22 (38)	19			21		
Maryland	17 (21)	13			14	20	(18)
Massachusetts	16 (15)	13	28 *	(28)	15	20	(17)
Michigan	17 (21)	14	18	(10)	18	13 *	(4)
Minnesota	10 (1)	11	6 *	(3)	9		
Mississippi	25 (47)	24 *	29	(11)	25	22	(6)
Missouri	14 (8)	12	18 *	(15)	14		
Montana	16 (15)	11	35	(32)	16	32 *	(28)
Nebraska	14 (8)	16	8 *	(1)	14		
Nevada	15 (12)	13			28 *		
New Hampshire	14 (8)	13			15		
New Jersey	20 (33)	16	25	(19)	16	24	(22)
New Mexico	22 (38)	21	24	(8)		23	
New York	12 (4)	8	22	(27)	7	18	(24)
North Carolina	19 (31)	11	41	(35)	15	24	(23)
North Dakota	11 (3)	10	17	(17)	11	17 *	(18)
Ohio	19 (31)	16	45	(34)	16	42	(30)
Oklahoma	17 (21)	13	18	(11)	17	23	(18)
Oregon	23 (40)	18			26		
Pennsylvania	16 (15)	15	22	(17)	14	26	(26)
Rhode Island	10 (1)	13 *			10		
South Carolina	23 (40)	20	25	(11)	25	21	(5)
South Dakota	16 (15)	14	17	(8)	15		
Tennessee	25 (47)	19	30	(24)	31	19	(1)
Texas	18 (28)	19	21	(7)	19	18	(7)
Utah	16 (15)	18			16		
Vermont	17 (21)	12			16		
Virginia	21 (36)	14	20	(15)	24	24	(10)
Washington	24 (43)	23	32	(19)	24	28	(15)
West Virginia	18 (28)	22	16	(2)	17		
Wisconsin	17 (21)	16	27	(24)	18	25 *	(21)
Wyoming	15 (12)	15			14		

* Interpret with caution

continued from page 7



Source: Sanders, William L. and Rivers, Joan C. "Cummulative And Residual Effects of Teachers on Future Student Academic Achievement," 1996, Table 1, p.

dents who attend high-poverty secondary schools. The same is true in Pennsylvania and Oklahoma: students in high-minority secondary schools are more likely to be taught by teachers without a college major in the subject they are teaching.

The problems in central cities are particularly acute, according to a 1995 report from the National Governors Association. "Emergency hiring, assignment of teachers outside their fields of preparation, and high turnover in underfunded schools conspire to produce a situation in which many poor and minority students are taught throughout their entire school careers by a steady stream of the least qualified and experienced teachers." [11]

A MORE EQUITABLE DISTRIBUTION OF TEACHER EXPERTISE

What would happen if minority and poor children had teachers of the same quality as other children? A large part of the gap would simply disappear. The estimates vary somewhat depending upon the statistical model used, but in no case is the effect minor.

- Ferguson's modeling for several metropolitan Alabama districts suggests that an increase of 1 standard deviation in the test scores of teachers who teach black children would produce a decline of about two-thirds in the black/white test score gap in that state.[12]
- Strauss's study of student achievement in North Carolina suggested that a 1% relative increase in teacher scores on the NTE would bring about a 5% relative decline in the percentage of students who fail standardized competency exams.[13]

In other words, much of what we have blamed on children and their families for decades is actually the result of things we have done to them. As a nation, we have deprived our neediest students of the very ingredient most important to learning: a highly qualified teacher.

In his analyses of the Texas data base, Ferguson found a small number of school districts that are exceptions to the general pattern (see below chart). A look at how their youngsters benefit from a steady diet of higher performing teachers gives us a glimpse of how the national data for poor and minority students *could* look...if we had the will.

ASSURING QUALIFIED TEACHERS FOR ALL OF OUR CHILDREN

These findings have profound implications for states and communities that are striving to get vastly larger numbers of their students to high standards of achievement. If education leaders want to accomplish this goal in the near term, they are far more likely to do so if they focus, first and foremost, on quality—quality in

1-68

teacher preparation, recruitment, hiring, assignment, and ongoing professional development.

This goes doubly for schools and communities with concentrations of poor and minority children. Rather than continuing to accept the crumbs, these schools and communities must insist on the very best teachers for their children. After all, poor and minority children depend on their teachers like no others. In the hands of our best teachers, the effects of poverty and institutional racism melt away, allowing these students to soar to the same heights as young Americans from more advantaged homes. But if they remain in the hands of underqualified teachers, poor and minority students will continue to fulfill society's limited expectations of them.

What, then, are the elements of a strategy to assure highly qualified teachers for all young Americans?

We don't yet have all the answers. But we know enough to start the conversations. Here are the more powerful ideas we have gleaned from our work with leading states and cities:

1. Standards for entry into the profession.

A number of states are raising the standards for entry into the profession. Virginia, for example, has raised both course requirements in the arts and sciences and cut scores on the Praxis examinations for aspiring teachers. Massachusetts has devised new and much more rigorous examinations, especially in the content areas.

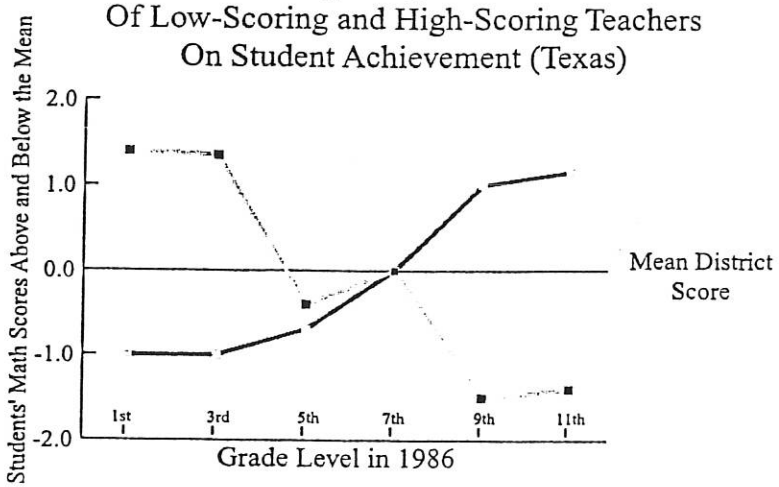
While these attempts are commendable, it is also important to make sure that the measures for teacher content knowledge are solid and aligned with K-12 standards. Preliminary information suggests that existing examinations may be too low: an analysis of a widely used test for prospective high school physics teachers, for example, featured content that one reviewer described as "appropriate for a rigorous ninth-grade physical science course." If this is correct, these tests are wholly insufficient either to assure adequate content knowledge of individual teachers or to use for accountability purposes with arts and sciences departments.

Any discussion about raising entry standards for teachers should include an examination of how well the standards align with the K-12 content candidates will have to teach, and the assessments used to find out if candidates can teach this content.

2. Accountability measures for colleges and universities that prepare teachers.

In Texas, for example, colleges that have pass rates below 70% (soon to be 75%) on the state's teacher licensure exam will lose the right to prepare teachers. To be sure that its intentions are understood, the legislature spells out precisely what it means: 70% of the

Long-Range Effects
Of Low-Scoring and High-Scoring Teachers
On Student Achievement (Texas)



— — Districts with Low-Scoring 1st and 3rd Graders and High-Scoring Teachers

▪ Districts with High Scoring 1st and 3rd Graders and Low-Scoring Teachers

Source: Ronald F. Ferguson, "Evidence That Schools Can Narrow the Black-White Test Score Gap," 1997.

white graduates, 70% of the Latino graduates, 70% of the black graduates and so on. Not a single group can be left behind. Moreover, if aspiring math teachers, for example, cannot pass the exam, then the math department loses the franchise. Other states are heading in this direction, as well. Universities, together with their nearby school districts, could take the lead from such state-level actions: decide on what intending teachers need to know in their subjects and hold academic departments accountable for getting them there before they graduate.

If education leaders want to close the achievement gap, they must focus, first and foremost, on developing qualified teachers.

3. Professional development for existing teachers.

Teacher effectiveness is not forever fixed. Through careful development, teachers can build their effectiveness over time. In Community School District #2 in New York City, Superintendent Tony Alvarado has invested generously in the professional development of his principals and teachers. Focusing initially on reading, and then moving to mathematics, Alvarado made sure his teachers, in particular, got lots of on-site coaching from experts. As a result, student achievement has climbed steadily over the past 10 years. University of Michigan researcher David Cohen's recent study of professional development in California also shows its impact on student achievement when professional development focuses on new curricula and the content that undergirds it.[14] Similar results are evident in broad achievement gains in the three El Paso school districts, where more than 50 full-time teacher-coaches provide in-school assistance to teachers as they strive to improve student achievement.

These successful strategies differ in important ways from many professional development programs and initiatives. Far from the three-hour workshop about isolated topics, these strategies are ongoing, on-site and focused on the content that students should learn.

4. Assurance that poor and minority children have teachers that are at least as qualified as the ones that teach other students.

Actually, if we had our druthers, we would push for a policy requiring that, for the next two decades or so, these students should systematically be assigned our best teachers. Achieving either goal, though, would require careful attention to:

- Just who we are preparing to teach—where they come from and where they want to teach, in particular;
- Interdistrict differences in salaries for beginning and mid-career teachers;
- The practice of concentrating beginning teachers in school buildings with concentrations of poor children;
- District policies—often gained through collective bargaining—that reward senior teachers with the "right" to transfer to "easier" schools;
- Practices within schools, where teachers fight over who has to teach whom, with the senior, better educated teachers often winding up with the most advanced children; and
- The absence of clear incentives and prevalence of disincentives for teachers to work with poor and minority children.

These practices have been around for so long that they seem beyond change. But some school districts are beginning to make headway on rooting out these inequities. In San Antonio, for example, new policies on teacher assignment have begun to balance the distribution of teachers within the district. In other districts, special targeting of more highly compensated "mentor" positions is beginning to even out teacher expertise. Energetic principals can also reverse the normal pattern. For example, in the Los Angeles unified School District, where uncertified and out-of-field teachers are the norm, Principal Lupe Simpson of the all-minority Nimitz Middle School has a mathematics department full of fully certified, mathematics majors. How? By working her contacts with local universities.

5. "Parent Right to Know" policies.

Parents deserve to know when their children are being taught science by history majors or history by physical education grads. To be sure, this knowledge has been available to some, mostly affluent parents through their community grapevines. But nowhere has there been a systematic way of letting all parents know that their child's teacher has enough background in the subject to teach it so their students will understand it. When parents know where the needs are greatest, they can become partners in local efforts to secure an adequate number of well-qualified teachers for all their students.

6. Recruitment and rewards to attract the best into teaching.

We worry that, instead of seeking out the very best, too many teacher preparation programs simply make do with what walks in the door. That's not good, because SAT and other data suggest that the high school seniors who aspire to become teachers are among the least able of all prospective college students. It's also not good for communities with concentrations of minority and poor students because few of those who aspire to become teachers either grew up in or want to teach in such communities.

Many leaders in teacher preparation programs say that they're doing the best they can—that low salaries and lower prestige make it impossible to attract able candidates, especially minorities, to the teaching profession and higher standards will make it worse. We remain unconvinced. If these claims are correct, then why does Teach for America, which has far higher standards than most education schools, routinely attract far more qualified graduates than it can place? And why, among Teach for America's way-above-average corps members, are there more than twice as many minorities as there are in education schools?[15] The same would appear to be true for alternate certification programs that cater to young or mid-career professionals from other fields: no lack of smart or minority applicants.

These experiences and others tell us that we can produce the highly qualified teachers that we need by combining:

- High entry standards;
- Rich incentives like generous scholarships and loan forgiveness for highly able professionals who want to teach in high-poverty schools;
- Accountability systems that reward departments and campuses for the numbers of their top students that enter teaching; and
- Non-traditional, yet still rigorous, routes into the profession.

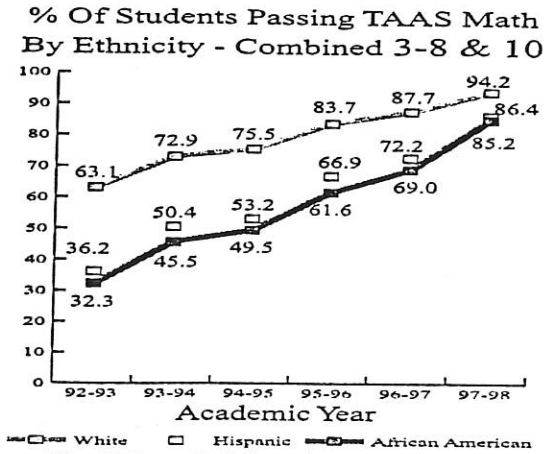
These are just some of the pieces of a solution to the vexing problem of assuring that we have teachers to match our goals. Solving this problem requires concerted action from policymakers, leaders in both K-12 and higher education, teacher unions, and parents. No single party can win the battle alone. All must be involved and at the table if we are to craft sound policies that will succeed.

But we must also understand that we cannot wait until every piece of this puzzle is in hand. Our inability to answer every question about teacher effectiveness right now shouldn't make us reluctant to use the devices we do have to begin to lure the best in, screen others out, and intensively develop the rest. And it certainly shouldn't deter us from doing what it takes to assure that poor and minority youngsters get at least their fair share of effective, well-prepared teachers.

We cannot wait until every piece of this puzzle is in hand. We must use the devices we have to lure the best teacher candidates in, screen others out, and develop the rest.

El Paso Closing the Gap

In 1992, leaders at the University of Texas-El Paso and the three El Paso-area school districts—El Paso, Ysleta, and Socorro—came together to create the El Paso Collaborative, a comprehensive effort to raise student achievement kindergarten through college. Their goal was to prepare every young person in this highly impoverished border city to be able to enter college without remediation, and the El Paso Standards they set reflected that goal.



The El Paso Collaborative for Academic Excellence, 1998.

Over the next five years, they focused hard on what matters most: excellent teaching. Through the Collaborative, El Paso teachers received intensive assistance in improving instruction, including summer institutes and regular on-site coaching, funded through a combination of NSF dollars and a redirection of federal and state funds. Meanwhile, leaders at the University made major changes in the way they were preparing teachers, to make sure that such teachers were fully prepared to teach to the El Paso standards.

The results of their hard work are clear in the data above: improved achievement and a narrower gap between groups. This is a refreshing change from the national picture of flat achievement and a widening gap between groups. Investing in teachers really does pay dividends!

NOTES:

1. Eric A. Hanushek. "The Trade-Off Between Child Quantity and Quality," *Journal of Political Economy*, 1992
2. William l. Sanders and June C. Rivers. "Cumulative and Residual Effects of Teachers on Future Students Academic Achievement," 1998, p. 9
3. Jeff Archer. "Students' Fortune Rests With Assigned Teacher," *Education Week*, Washington, DC, February 18, 1998.
4. Ronald F. Ferguson. "Evidence That Schools Can Narrow the Black-White Test Score Gap," 1997, p. 32
5. Ronald F. Ferguson and Helen F. Ladd. "How and Why Money Matters: An Analysis of Alabama Schools," in *Holding Schools Accountable: Performance Based Reform in Education*, Brookings Institute: Washington, DC, 1996
6. Dan D. Goldhaber and Dominic J. Brewer, "Evaluating the Effect of Teacher Degree Level on Educational Performance," in *Developments in School Finance*, 1996, p. 199
7. Eva L. Baker, "Report on the Content Area Performance Assessments (CAPA): A Collaboration Among Hawaii Dept. of Education, the Center for Research on Evaluation Standards and Student Testing, and the Teachers and Children of Hawaii" 1996, p.17
8. John F. Kain and Kraig Singleton. "Equality of Educational Opportunity Revisited," *New England Economic Review*, May/June 1996, p. 109
9. Ferguson, "Evidence That Schools Can Narrow the Black-White Test Score Gap," p. 30
10. Kain and Singleton, p. 109
11. Linda Darling-Hammond, "The Role of Teacher Expertise and Experience in Students' Opportunity to Learn," in *Strategies for Linking School Finance and Students' Opportunity to Learn*, National Governors Association, Washington DC: 1996
13. Ferguson and Ladd, p. 278
14. Robert P. Strauss and Elizabeth A. Sawyer. "Some New Evidence on Teacher and Student Competencies," *Economics of Education Review*, 1986, p.41
15. David K. Cohen and Heather C. Hill. "State Policy and Classroom Performance," *CPRE Policy Briefs*, January 1998
16. Teach for America, 1997 Annual Report, New York NY

A MATTER OF QUALITY

A Strategy for Assuring the High Caliber of America's Teachers

by Lowell Milken

A MATTER OF QUALITY
A STRATEGY FOR ASSURING THE HIGH CALIBER OF AMERICA'S TEACHERS

SUMMARY

"Good teachers are to education what education is to all other professions—the indispensable element, the sunlight and oxygen, the foundation on which everything else is built. They are central to assuring excellence and rigor in the educational experience of every young person in America. Yet our present K-12 system is *not* providing children with high-caliber teachers every year they are in school.

We therefore propose a new strategy for the American teaching profession. It neither abandons our public education system nor tinkers at the margin. Instead, it honors its *essence* yet changes the structure of the profession. For it is my belief that, unless we take bold and immediate action, we will not attract, retain and motivate the high quality educators we need to provide *all* children with the education they need and deserve.

After years of closing its eyes to the role of teaching, the nation is beginning to look squarely at the importance of teacher quality. Nine out of ten Americans say that high quality educators are second only to student safety as the most important issue facing education—and recent academic research confirms their view of the importance of teachers in student learning.

The need to strengthen K-12 education is also an urgent economic imperative. Where earlier eras required physical and financial capital to fuel the agricultural and industrial sectors, our era requires human capital to fuel "knowledge industries"—specifically, those propelled by information and communication technology. Education, training and skills now account for about 75 percent of the nation's wealth and growth. Our knowledge-based economy is creating tremendous competition for strong human resources. Already U.S. businesses cannot find enough skilled workers."

"Third. An *expanded pool of high quality teachers* is ensured by a comprehensive approach to enlisting the best. An initial academic degree and teaching certification are attainable in four years or through rigorous assessments, thus creating multiple entry paths. A mandated salaried, mentored induction year gives new teachers classroom responsibility with intensive support. All teachers have the opportunity to become nationally certified. Proactive national recruitment strategies identify additional new sources of talent, through Educational Leadership Scholarships, incentives for mid-career professionals interested in teaching and a national advertising campaign that conveys a new, positive and attainable vision of what it means to be a teacher."

Journal of Personnel Evaluation in Education 11: 57–67, 1997

Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation

S. PAUL WRIGHT, SANDRA P. HORN AND WILLIAM L. SANDERS

University of Tennessee, Value-Added Research and Assessment Center, 225 Morgan Hall, P.O. Box 1071, Knoxville, Tennessee 37901-1071

Abstract

The Tennessee Value-Added Assessment System (TVAAS) has been designed to use statistical mixed-model methodologies to conduct multivariate, longitudinal analyses of student achievement to make estimates of school, class size, teacher, and other effects. This study examined the relative magnitude of teacher effects on student achievement while simultaneously considering the influences of intraclassroom heterogeneity, student achievement level, and class size on academic growth. The results show that teacher effects are dominant factors affecting student academic gain and that the classroom context variables of heterogeneity among students and class sizes have relatively little influence on academic gain. Thus, a major conclusion is that teachers make a difference. Implications of the findings for teacher evaluation and future research are discussed.

Overview

Over the years, educational researchers have investigated many factors considered to affect student learning. At the heart of this line of inquiry is the core belief that *teachers make a difference*. There are continuing debates about how much the extant teacher-effectiveness literature (e.g., Brophy, 1986; Porter & Brophy, 1988) can be trusted to identify characteristics of effective teachers, and additional debates as well about how such research findings should frame the subsequent development of teacher evaluation systems (e.g., Ellett, 1990; Scriven, 1990; Peterson, Kromrey & Smith, 1990). In addition, there is considerable argument over the logic behind and the extent to which student achievement data should be used as a basis for teacher evaluation (Berk, 1988; Schalock & Schalock, 1993). These debates aside, few attempts have been made to directly measure the influence of individual teachers on the academic progress of large *populations of students* using measurements available from traditional standardized testing programs. Partial confounding of educational (teacher) effects with factors exogenous to schooling influences (see Wang, Haertel & Walberg, 1993 for an explication of these issues) and the nonrandom assignment of students to teachers are two of the reasons most often assumed to be insurmountable obstacles to this type of inquiry.

In criticizing and arguing equity issues in the fair application of teacher evaluation instruments and procedures, teachers have often directed their comments to classroom context characteristics. Key among these has been the issue of the ability level of students and the range in individual differences among students in ability levels. As the argument

typically proceeds, teachers who have classes more heterogeneous than homogeneous in ability levels are at a distinct disadvantage in producing effects on student learning and subsequent achievement, particularly as inferred from standardized test scores.

Recently, new processes for estimating the effects of teachers and schools on student academic outcomes free of these traditional objections have been developed. One of these—the Tennessee Value-Added Assessment System (TVAAS), which uses statistical mixed-model methodology to enable a multivariate, longitudinal analysis of student achievement data— has been demonstrated to produce estimates of school and teacher effects that are free of socioeconomic confoundings and do not require direct measures of these concomitant variables (see Sanders & Horn, 1995b, and Sanders, Saxton & Horn, in press, for greater detail). To support TVAAS, a massive database of longitudinally merged student, teacher, school, and school system information has been compiled for the primary purpose of determining system, school, and teacher effects on the academic gains of students. Utilizing this database, the present study attempts to measure the relative magnitude of teacher effects while simultaneously considering the influences of intraclassroom heterogeneity, student achievement level, and class size on academic growth. Among these influences, intraclassroom heterogeneity is of special interest. The magnitude of this variability may be a natural occurrence or can result from intentional grouping of students. Regardless of cause, the evaluation of the influence of intraclassroom, variability on the academic growth of student populations and its interaction with teacher effects is another important research objective of this study.

Methodology

For the purpose of this investigation, results are derived from analyses of a subset of data from the 1994 and 1995 TCAP scores for five subjects (math total, reading total, language total, social studies, and science) and three grades (third, fourth, and fifth). TCAP tests are given each spring to all students in Tennessee in grades two through eight. An important property of these tests is that the scale scores form a single, continuous, equal-interval scale across all grades (CTB/McGraw-Hill, 1990, pp. 4–5), allowing for measurement of student academic progress from year to year. The analyses reported here are based on student academic gain—that is, the student's scale score this year minus that student's scale score last year. Thirty separate analyses were done. Each of the fifteen subject–grade combinations was analyzed separately, and each of these fifteen analyses was carried out on two different sets of school systems in Tennessee. One set consisted of thirty East Tennessee school systems, and the other consisted of twenty-four Middle Tennessee systems. A mixed-model analysis of variance was obtained by fitting the following model[1] to the data:

$$Y = M + S + H + C + H^*C + T(S^*H^*C) + A + A^*S + A^*H + A^*C + A^*H^*C + A^*T(S^*H) + E,$$

where

Y is the student's gain score,
M is an overall mean gain,
S is the school system,
H is heterogeneity-in-achievement (three groups were used),
C is the class size (two groups were used),
$H*C$ is the heterogeneity-by-class-size interaction,
$T(S*H*C)$ is the teacher, each one nested within a particular combination of system, heterogeneity groups, and class-size group,
A is achievement level (four groups were used),
$A*S$ is the achievement-by-system interaction,
$A*H$ is the achievement-by-heterogeneity interaction,
$A*C$ is the achievement-by-class-size interaction,
$A*H*C$ is the achievement-by-heterogeneity-by-class-size interaction,
$A*T(S*H*C)$ is the achievement-by-teacher interaction,
E is the random "error" term.

The $T(S*H*C)$, $A*T(S*H*C)$, and E terms represent random effects. All the other effects are fixed. The analyses were done with the MIXED procedure in SAS/STAT version 6.09 running on an IBM RS/6000 Model 590 work station at the Value-Added Research and Assessment Center at the University of Tennessee, Knoxville.

The response variable—the educational outcome of the student—was the student's gain score from 1994 to 1995—that is, the student's 1995 scale score on the TCAP minus the student's 1994 scale score. The student's achievement level was defined operationally as the average of the student's 1994 and 1995 scale scores. Classroom heterogeneity in achievement was defined operationally as the standard deviation of the achievement level scores of the students in the class, as defined above. The larger the standard deviation, the more heterogeneous in achievement were the students in the class. For the analysis, classrooms were classified into three groups—low, moderate, and high heterogeneity—using their standard deviation of achievement level. The moderate group contained about half of the classrooms, and the two extreme groups each contained about one-fourth of the classrooms. Students were classified into four achievement level groups of roughly equal size using the achievement level scores described above. Inclusion of an achievement level variable was thought to be particularly important in view of the results of earlier studies indicating that the value of tracking or not tracking depended on the achievement level of the student (Kulik, 1992).

Two class-size groups were used: small (ten to nineteen students) and large (twenty to thirty-two students). Classes of fewer than ten or more than thirty-two students were omitted. There were several reasons for omitting the larger classes. The first was that the database currently does not actually identify the classroom of each student. It does identify the teacher for each student and subject. The reason that only third, fourth, and fifth grades were analyzed is because, in these grades, it is more commonly the case that each student is in a single classroom with a single teacher. Nevertheless, some teachers in the database

were shown to have a large number of students, too many to represent a single classroom. Omitting teachers with more than thirty-two students provided a way to avoid treating as one classroom what was in fact several classes taught by the same teacher.

Results

Table 1 through 3 summarize the results for grades three through five, respectively. As an aid for assessing both the statistical significance and the effect sizes of the various effects in the model, z-scores are reported for each effect. For random effects, z-scores were obtained by dividing the estimated variance component for the effect by its estimated standard error. For large samples (such as those in this study), this z-score is approximately distributed as a standard normal variate. For fixed effects, first p-values were obtained

Table 1. z-Values for Analyses of Third-Grade Gains.

Source	*Set*	*Math*	*Reading*	*Language*	*Social Studies*	*Science*
System (S)	1	6.12	2.26	4.34	4.03	3.13
	2	4.86	3.55	5.39	5.55	3.92
Heterogeneity (H)	1	1.39	0.25	0.61	0.81	0.05
	2	1.54	0.09	1.64	0.61	0.30
Class size (C)	1	0.57	0.02	1.45	0.14	1.92
	2	1.03	0.64	0.16	0.97	0.38
H*C	1	0.58	0.49	0.29	0.45	1.83
	2	0.20	0.47	2.21	0.20	0.83
Teacher (S*H*C) (T)	1	12.48	7.85	11.04	6.09	7.76
	2	13.14	8.69	12.06	8.33	8.88
Achievement level (A)	1	17.00	12.65	8.49	10.04	6.76
	2	28.04	20.14	8.96	14.53	8.41
A*S	1	2.19	1.88	2.70	2.49	2.19
	2	1.25	5.31	1.46	3.34	3.26
A*H	1	2.05	4.64	1.15	4.36	0.53
	2	1.41	0.76	1.29	3.78	4.27
A*C	1	1.37	0.53	0.40	0.18	1.53
	2	0.12	0.67	1.14	2.33	1.19
A*H*C	1	0.07	0.22	0.32	0.10	0.70
	2	2.05	0.94	0.37	2.12	2.18
A*T	1	2.35	4.88	2.02	0.61	1.05
	2	0.73	0.68	1.27	1.69	2.39
N	1	10751	10564	10916	10005	9939
	2	13632	13506	14079	13651	13624

Set: 1 = 30 East Tennessee school systems.
2 = 24 Middle Tennessee school systems.
N = total number of students.

from F statistics, then corresponding z-scores were calculated from the p-values by treating the p-values as if they were two-tailed and from a standard normal distribution. This technique of converting p-values to z-scores is commonly used in meta-analysis to convert results from a variety of tests to a common metric (see, for example, Rosenthal, 1984, p. 65). For reference, the z-values correspond to the two-tailed p-values of 0.10, 0.05, 0.01, 0.001, and 0.0001 are 1.64, 1.96, 2.58, 3.29, and 3.89, respectively.

It is clear from Tables 1 to 3 that the two most important factors impacting student gain are the teacher and the achievement level for the student. The teacher effect is highly significant in every analysis and has a larger effect size than any other factor in twenty of the thirty analyses. The achievement-level effect is significant in twenty-six of the thirty analyses and has the largest effect size in ten of the thirty analyses. These results are discussed in more detail in the Discussion section below.

The third most important factor overall was the school system. There were significant

Table 2. z-Values for Analyses of Fourth-Grade Gains.

Source	*Set*	*Math*	*Reading*	*Language*	*Social Studies*	*Science*
System (*S*)	1	5.63	3.66	5.68	4.23	2.55
	2	5.56	5.07	4.62	4.02	3.00
Heterogeneity (*H*)	1	0.20	0.03	0.13	2.53	0.62
	2	1.84	1.32	0.94	1.47	1.00
Class size (*C*)	1	1.65	1.00	1.30	2.83	1.47
	2	0.39	1.14	1.14	0.81	0.49
H*C	1	2.29	0.80	0.98	2.30	0.75
	2	1.31	0.69	0.62	2.40	1.11
Teacher (*S*H*C*) (*T*)	1	11.17	6.04	9.24	7.17	7.93
	2	12.49	5.72	10.48	6.69	7.62
Achievement level (*A*)	1	2.45	13.04	8.61	3.37	10.99
	2	6.70	11.92	8.36	4.59	10.91
*A*S*	1	2.63	3.01	1.86	2.14	1.55
	2	3.50	4.50	1.43	5.27	3.74
*A*H*	1	0.28	1.32	2.53	2.01	0.12
	2	0.59	0.89	1.02	0.55	2.06
*A*C*	1	2.96	0.84	1.18	1.53	0.34
	2	1.09	1.99	0.99	0.42	1.68
*A*H*C*	1	1.13	1.33	0.02	0.73	1.25
	2	1.50	0.18	0.05	1.09	0.78
*A*T*	1	1.75	0.56	1.40	2.45	1.24
	2	2.14	2.61	1.10	1.06	0.47
N	1	10344	10477	10497	9438	9329
	2	13102	13102	13498	12320	12406

Set: 1 = 30 East Tennessee school systems.
2 = 24 Middle Tennessee school systems.
N = total number of students.

differences among school systems in twenty-seven of the thirty analyses, and the effect sizes are in most cases impressively large, though not nearly as large as for the teacher and achievement-level factors. A notably nonsignificant factor was class size. The main effect for class size was significant in only three of the thirty analyses. In two of these three instances, the smaller-size class had the higher gains; in the other case, the larger-size class had higher gains. Class size also appeared in a number of statistically significant interactions, though most of these had relatively small effect sizes. The interpretations of these interactions are as varied as those for the class-size main effect. Since the objective was not to investigate the class size effect per se but merely to control for that effect where it occurs, no further discussion of this point is offered.

Based upon an effect size (z-value) of 2.0 (corresponding to a significance level of approximately 0.05), the main effect for heterogeneity was statistically significant in only two of the thirty analyses, approximately the number that would be expected to occur by

Table 3. z-Values for Analyses of Fifth-Grade Gains.

Source	*Set*	*Math*	*Reading*	*Language*	*Social Studies*	*Science*
System (*S*)	1	1.30	3.52	3.18	1.04	1.30
	2	5.69	3.50	2.49	4.20	3.02
Heterogeneity (*H*)	1	0.55	0.57	1.44	0.37	2.56
	2	0.66	0.33	1.41	0.12	0.59
Class size (*C*)	1	2.19	0.72	0.59	1.58	2.35
	2	1.13	1.40	0.71	0.14	0.01
H*C	1	0.29	0.82	0.23	1.13	1.77
	2	0.66	0.79	1.37	0.10	0.11
Teacher (*S*H*C*) (*T*)	1	9.70	5.80	6.29	5.65	6.24
	2	9.13	6.33	9.68	6.62	6.27
Achievement level (*A*)	1	1.94	4.42	1.51	0.14	5.20
	2	3.88	5.12	2.26	1.29	2.24
*A*S*	1	2.60	2.03	2.64	0.91	2.15
	2	3.36	2.15	0.98	4.24	0.59
*A*H*	1	2.81	1.07	1.10	0.78	1.18
	2	0.70	2.40	0.91	1.22	0.97
*A*C*	1	2.07	1.09	1.70	0.94	0.93
	2	2.35	1.18	0.13	0.86	0.88
*A*H*C*	1	1.49	0.06	1.31	0.24	1.63
	2	1.46	0.39	1.43	0.45	3.04
*A*T*	1	1.79	2.52	1.52	0.05	0.63
	2	3.48	0.64	0.00	0.00	1.87
N	1	8259	8874	8615	6527	6662
	2	9939	9629	10141	9136	8569

Set: 1 = 30 East Tennessee school systems.
2 = 24 Middle Tennessee school systems.
N = total number of students.

chance. The statistically significant effects for heterogeneity were found in fourth-grade social studies and fifth-grade science in East Tennessee. In the first instance, the estimated mean gains for the three groups (low, moderate, and high heterogeneity) were 26.9, 26.4, and 21.6. In the second instance, the estimated mean gains were 10.8, 10.7, and 15.9. So in one case, higher gains occurred under lower heterogeneity, and in the other case higher gains occurred under higher heterogeneity. (Note that the scales for social studies and science are not comparable, so the larger point gains in social studies do not indicate greater academic progress than the smaller ones indicated for science.)

In addition to significant main effects, there were a number of statistically significant interactions, including a significant three-way interaction of achievement level, heterogeneity, and class size in four of the thirty analyses. Specifically, in the thirty analyses there were a total of 180 interaction effects of which fifty-one were statistically significant. However, the effect sizes were relatively small: only seventeen exceeded 3.0 (in absolute value) and only eight exceeded 4.0. The largest interaction effect had a z-value of 5.31. For comparison, the smallest teacher effect size was 5.65. While some of the interaction effects appear to be different from zero, their interpretation tends to vary from subject to subject and grade to grade so that no general conclusions can be drawn. For example, there were seventeen significant interactions involving the heterogeneity factor (out of a total of ninety interactions involving heterogeneity in the thirty analyses), mostly with relatively small effect sizes. From these analyses, we conclude that the effect of intraclassroom heterogeneity neither as a main effect nor interacting with other factors is important in the academic growth of students.

Discussion

Despite ongoing debates about whether, and how much teachers make a difference in student learning relative to a host of other factors assumedly affecting student learning (Wang, Haertel & Walberg, 1993), and whether particular elements of teaching can be systematically and causally linked to student achievement (Scriven, 1990), the results of this study well document that the most important factor affecting student learning is the teacher. In addition, the results show wide variation in effectiveness among teachers. The immediate and clear implication of this finding is that seemingly more can be done to improve education by improving the effectiveness of teachers than by any other single factor. *Effective teachers appear to be effective with students of all achievement levels, regardless of the level of heterogeneity in their classrooms.* If the teacher is ineffective, students under that teacher's tutelage will achieve inadequate progress academically, regardless of how similar or different they are regarding their academic achievement. This finding is corroborated by recent research on the cumulative effects of teachers on the academic progress of students (Sanders & Rivers, 1996). These recent studies show that teacher effects on student learning as inferred from standardized test scores are additive and cumulative over grade levels with little evidence of compensatory effects. Thus, students in classrooms of very effective teachers, following relatively ineffective teachers,

make excellent academic gains but not enough to offset previous evidence of less than expected gains.

The other dominant factor in the results of the analyses reported here was the achievement level of the student. Table 4 shows the estimated mean gains in each achievement

Table 4. Estimated Mean Gains by Four Achievement Levels with Standard Errors in Parentheses.

		Achievement Level				
	Set	*Lowest*			*Highest*	*z*
Third grade	1	64.2 (1.6)	56.0 (1.4)	45.2 (1.4)	35.9 (1.4)	17.0
	2	75.4 (1.2)	59.3 (1.2)	47.5 (1.1)	36.6 (1.1)	28.0
Fourth grade	1	20.8 (1.4)	19.3 (1.1)	19.9 (1.1)	16.1 (1.2)	2.5
	2	28.7 (1.1)	25.7 (1.1)	21.4 (1.0)	20.5 (1.0)	6.7
Fifth grade	1	23.6 (1.4)	26.1 (1.2)	27.0 (1.2)	24.0 (1.3)	1.9
	2	25.9 (1.1)	27.2 (1.0)	25.9 (1.1)	21.2 (1.2)	3.9
Reading:						
Third grade	1	42.5 (1.5)	34.0 (1.2)	27.7 (1.3)	19.4 (1.3)	12.7
	2	45.3 (1.2)	33.0 (1.0)	26.6 (1.0)	16.4 (1.0)	20.1
Fourth grade	1	10.5 (1.1)	16.8 (0.9)	20.4 (1.0)	28.5 (1.0)	13.0
	2	16.7 (1.0)	20.8 (0.9)	22.9 (0.9)	32.6 (1.0)	11.9
Fifth grade	1	9.7 (1.3)	9.7 (1.1)	16.0 (1.1)	13.6 (1.1)	4.4
	2	11.6 (1.1)	10.3 (1.1)	16.0 (1.0)	17.4 (1.1)	5.1
Language:						
Third grade	1	29.7 (1.1)	25.1 (1.0)	18.4 (1.0)	23.0 (1.0)	8.5
		30.7 (0.9)	26.6 (0.8)	21.3 (0.8)	23.4 (0.8)	9.0
Fourth grade	1	10.7 (1.1)	20.0 (1.0)	18.5 (1.0)	23.4 (1.1)	8.6
	2	16.2 (1.0)	21.7 (1.0)	21.1 (0.9)	27.3 (1.0)	8.4
Fifth grade	1	14.8 (1.1)	16.9 (1.1)	15.8 (1.0)	17.9 (1.1)	1.5
	2	13.5 (1.0)	14.6 (1.1)	15.8 (1.0)	17.5 (1.1)	2.3
Social studies:						
Third grade	1	40.8 (2.0)	46.9 (1.7)	37.1 (1.6)	24.4 (1.6)	10.0
	2	46.2 (1.7)	49.0 (1.4)	39.8 (1.3)	23.6 (1.4)	14.5
Fourth grade	1	26.7 (1.9)	27.5 (1.6)	26.3 (1.6)	19.5 (1.7)	3.4
	2	28.5 (1.6)	31.4 (1.4)	29.4 (1.4)	22.3 (1.4)	4.6
Fifth grade	1	30.2 (1.8)	30.1 (1.6)	29.1 (1.6)	30.8 (1.8)	0.1
	2	28.9 (1.6)	28.3 (1.5)	25.6 (1.5)	25.7 (1.3)	1.3
Science:						
Third grade	1	18.1 (1.9)	28.5 (1.5)	24.5 (1.5)	15.9 (1.5)	6.8
	2	23.3 (1.5)	30.1 (1.3)	25.2 (1.2)	15.8 (1.3)	8.4
Fourth grade	1	24.9 (1.7)	22.6 (1.4)	17.6 (1.4)	5.6 (1.4)	11.0
	2	25.0 (1.5)	24.4 (1.2)	20.0 (1.2)	8.3 (1.3)	10.9
Fifth grade	1	19.6 (1.7)	10.2 (1.5)	8.2 (1.4)	11.8 (1.6)	5.2
	2	13.7 (1.6)	9.4 (1.4)	9.3 (1.3)	12.9 (1.3)	2.2

Set: 1 = 30 East Tennessee school systems.
2 = 24 Middle Tennessee school systems.

level group for all thirty analyses (including four in which the effect was not statistically significant). No universally applicable pattern emerges, but it is worth noting that out of the twenty-six analyses in which achievement level was significant, the largest gains occurred in the lowest achievement group twelve times, in one of the two middle groups eight times, and in the highest group six times. Similarly, the smallest gains occurred in the highest achievement group fifteen times, in one of the two middle groups six times, and in the lowest group five times. In other words, there is a disturbingly common but not universal pattern for the best students to make the lowest gains. Possible explanations include a lack of stretch in curriculum and instruction to accommodate the highest achievers and insufficient availability of higher level course offering in all schools.

Hundreds of studies on ability grouping have been conducted since the 1930s. Recent meta-analyses of these studies by Slavin (1987, 1990) and Kulik (1992) have synthesized the findings of the most rigorous studies. Slavin, in both of his studies, discovered that "study after study, including randomized experiments of a quality rarely seen in educational research, finds no positive effect of ability grouping in any subject or at any grade level, even for the high achievers most widely assumed to benefit from grouping" (Slavin, 1990, p. 491). Experts on ability grouping contend that the effects of grouping on achievement are minimal except in classrooms where there is significant curricular adjustment to meet the needs of students at different levels (Kulik, 1992; O'Neil, 1992; Rogers & Kimpston, 1992). Slavin (1990, p. 491) goes so far as to suggest that "the lesson to be drawn from research on ability grouping may be that unless teaching methods are systematically changed, school organization has little impact on student achievement," This study supports Slavin's conclusion.

Teachers seem to have far more to do with the academic progress of students than does the method used for assignment of children to teachers. The contention that high academic gains are more likely to be produced in highly homogeneous classrooms is not supported by our research, and, therefore, neither is the corollary that teachers with highly heterogeneous classrooms should not be expected to make those gains.

Perhaps the persistence of the phenomenon of ability grouping in American schools, despite the preponderance of research attesting to its ineffectiveness, can be attributed to the reluctance of the educational community to assign responsibility for student achievement to teachers. Travers (1981, p. 18) expresses this point of view thusly: "The extent to which a pupil learns in the school is a function of many different conditions, of which the teacher's mode of operation is only one. . . . The teacher factor may well account for only a small amount of the differences in achievement." Such statements as these, in turn, may derive from two widely held beliefs: that the interplay of the educational setting with factors outside the purview of formal education prevents the correct attribution of learning effects; and that most educational assessment tools and standardized tests, in particular, are poor indicators of academic progress (for a discussion of this latter point, see Sanders & Horn, 1995a). However, these beliefs do not seem supported and are contrary to the findings of this study. It is recognized here, however, that identifying a common set of factors and interpretation of their effects on student learning and achievement presents a highly complex set of methodological and theoretical issues (Wang, Haertel & Walberg, 1993).

Conclusions and Implications

Differences in teacher effectiveness were found to be the dominant factor affecting student academic gain. The importance of the effects of certain classroom contextual variables (class size and classroom heterogeneity) appears to be minor and should not be viewed as inhibitors to the appropriate use of student outcome data in teacher assessment. These results suggest that teacher evaluation processes should include, as a major component, a reliable and valid measure of a teacher's effect on student academic growth over time. The use of student achievement data from an appropriately drawn standardized testing program administered longitudinally and appropriately analyzed can fulfill these requirements. If the ultimate goal is to improve the academic growth of student populations, one must conclude that improvement of student learning begins with the improvement of relatively ineffective teachers regardless of the student placement strategies deployed within a school.

In addition, student academic level was found to be significantly related to academic progress, although not nearly to the degree found for the teacher. Disproportionately, high-scoring students were found to make somewhat lower gains than average and lower-scoring students. Possible explanations include lack of opportunity for high-scoring students to proceed at their own pace, lack of challenging materials, lack of accelerated course offerings, and concentration of instruction on the average or below-average student. This finding indicates that it cannot be assumed that higher-achieving students will "make it on their own."

Though the debate about whether student achievement data should be used as part of an assessment, evaluation, and accountability system for teachers will assuredly continue, the results of this study suggest that *teachers do make a difference* in student achievement. It is recognized here, however, that there were no direct, systematic observations of the quality of teaching and learning at the classroom level in this study. Thus, identifying teachers that clearly get results over time, and comparing them to teachers over time who do not, seems a logical, worthwhile next step in addressing the issues raised here and in further developing general lines of inquiry about the important relationship between teacher effectiveness and teacher evaluation. If characteristics of teaching and learning environments that differentiate teachers who are demonstrably effective (as opposed to ineffective) in different contexts over time can be documented, subsequent teacher evaluation systems might be developed to accommodate these characteristics. Continuing debates aside, the results presented here suggest that *teachers indeed make a difference* and that homogeneity and heterogeneity of student ability levels within classes are not major concerns in assessing teacher effectiveness. Those developing future teacher evaluation systems might take comfort in the results reported here with the suggestion that variation in ability levels of students, despite teacher arguments and conventional wisdom, is not a major factor framing effectiveness in teaching.

Notes

1. This model would not be adequate and appropriate to provide the best possible estimate of an individual effect. Rather the full TVAAS model should be used (Sanders, Saxton & Horn, in press).

References

Berk, R. (1988). Fifty reasons why student achievement gains does not mean teacher effectiveness. *Journal of Personnel Evaluation in Education*, 1(4) 345–364.

Brophy, J. (1968). Teacher influences on student achievement. *American Psychologist* (October), 1069–1077.

CTB/McGraw-Hill (1990). *Comprehensive test of basic skills* (4th ed.). Spring Norms Book. Monterey, CA: CTB/Macmillan/McGraw-Hill.

Ellett, C. D. (1990). *A new generation of classroom-based assessments of teaching and learning: Concepts, issues and controversies from pilots of the Louisiana STAR*. Baton Rouge: Teaching Internship and Statewide Teacher Evaluation Projects, College of Education, Louisiana State University.

Kulik, J. A. (1992). *An analysis of the research on ability grouping: Historical and contemporary perspectives*. National Research Center on the Gifted and Talented. Storrs, CT: University of Connecticut.

O'Neil, J. (1992). On tracking and individual differences: A conversation with Jeannie Oakes. *Educational Leadership*, 50(2), 18–21.

Peterson, D., Kromrey, J., & Smith, D. C. (1990). Research-based teacher evaluation: A response to Scriven. *Journal of Personnel Evaluation in Education*, 4(1), 7–18.

Porter, A. C., & Brophy, J. (1988). Synthesis of research on good teaching: Insights from the work of the Institute for Research on Teaching. *Educational Leadership*, 45(8), (May), 74–85.

Rogers, K. B., & Kimpston, R. D. (1992). Acceleration: What we do vs. what we know. *Educational Leadership*, 50(2), 58–61.

Rosenthal, R. (2984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Sanders, W. L., & Horn, S. P. (1995a). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Educational Policy Analysis Archives*, 3(6).

Sanders, W. L., & Horn, S. P. (1995b). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. In A. J. Shinkfield & D. Stufflebeam (Eds.), *Teacher evaluation: Guide to effective practice* (pp. 337–350). Boston: Kluwer.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Research Progress Report. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sanders, W. L., Saxton, A. M., & Horn, S. P. (In press). The Tennessee Value-Added Assessment Systems (TVAAS): A quantitative, outcome-based approach to educational assessment. In J. Millman (Ed.), *Assuring accountability? Using gains in student learning to evaluate teachers and schools*. Thousand Oaks, CA: Corwin Press.

Schalock, H. A., & Schalock, M. D. (1993). Student learning in teacher evaluation and school improvement: An introduction. *Journal of Personnel Evaluation in Education*, 7(2), 103–104.

Scriven, M. (1990). Can research-based teacher evaluation be saved? *Journal of Personnel Evaluation in Education*, 4(1), 19–39.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293–336.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research*, 60(3), 471–499.

Travers, R. M. W. (1981). Criteria of good teaching. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 14–22). Beverly Hills, CA: Sage.

Wang, M. C., Haertel, G., & Walberg, H. J. (1993). Toward a knowledge base of school learning. *Review of Educational Research*, 73(3), 249–294.

THOMAS B. FORDHAM FOUNDATION
OUTSIDE THE BOX

The Teachers We Need and How to Get More of Them

"Every child should be able to count on having a teacher who has a solid general education, who possesses deep subject area knowledge, and who has no record of misbehavior. The state has an obligation to ensure that all prospective teachers meet this minimal standard. Thus states should perform background checks on candidates for teaching positions. To boost the likelihood that those who teach our children are themselves well educated, states should require that teaching candidates have at least a bachelor's degree in some academic subject.

States should also ensure subject matter competence. There are two ways to do this: requiring teachers to major in the subjects they teach or requiring them to pass challenging tests of subject matter knowledge. Neither method is perfect. Obliging all teachers to major in the subject they will teach may-regrettably-set the bar too low. At some universities, one can graduate as a history major without learning much of the history we'd expect a high school history teacher to have mastered. The same is true of other academic majors. And a minor is unlikely to reflect any subject mastery. On the other hand, a prospective teacher who graduates in, say, American studies may have learned ample history or literature to be an outstanding history or English teacher, even though his diploma doesn't actually say "history" or "English".

Such variation in college majors tempts us to embrace testing as a more reliable measure of preparedness to teach. The value of any test, however, hinges on its content, rigor and passing score. Our instinct is to set those cut-offs as high as possible. But since tests are an imperfect gauge of teaching ability, some applicants will fail the test yet possess superior teaching potential.

We all know individuals whose other qualities would cause them to be effective with children even if they do poorly on a paper-and-pencil test of knowledge. That is why we are wary of putting all the education eggs in the testing basket or making a certain fixed score an absolute prerequisite to being hired.

Neither academic majors nor subject test scores is a faultless means of assuring that teachers possess the requisite knowledge and will be good at delivering it. But either strategy is superior to today's widespread disregard of subject-matter mastery."

(www.edexcellence.net/library/teacher.html)

"States should expand the pool of talented teaching candidates by allowing individuals

who have not attended schools of education to teach, provided that they meet the minimum standards outlined above. States should encourage programs that provide compressed basic training for prospective teachers. States should also attract outstanding college graduates to the profession by using financial incentives such as scholarships, loan forgiveness programs and signing bonuses."

(www.edexcellence.net/library/teacher.html)

EDUCATION COMMISSION OF THE STATES

ECS

The Progress of Education Reform 1998

WHAT'S INSIDE

Technology

The presence of up-to-date technology and telecommunications in America's schools continues to increase dramatically. Today, three out of four classrooms have at least one computer designated for instructional use, and 85% of the nation's schools are connected to the Internet.

But as public spending on school technology grows – an estimated $5 billion a year nationwide – so does the pressure on educators, schools and districts to demonstrate the investment is paying off.

"It is no longer enough for educators to simply report . . . a better student-to-computer ratio or an increase in the number of wired classrooms," noted a recent analysis by the Milken Exchange on Education Technology. "Policymakers want evidence that technology is being used to improve student learning."

Assessing the value of education technology, however, is easier said than done. Most of the research done so far has been anecdotal, small-scale and, in many cases, based on measures not everyone agrees are the right ones to use – namely, standardized test scores. Such tests do not reflect the full range of benefits students can get from using technology.

The Milken report urges a broader, more balanced approach. It calls for a national research agenda focused on identifying the "essential conditions" for maximizing the potential of technology as a teaching and learning tool. It also offers a set of indicators states can use, in the interim, to help assess their progress.

Technology in American Schools: Seven Dimensions for Gauging Progress provides a framework of progress indicators to help states chart their course toward effective use of technology in schools. It is available online (www.milkenexchange.org) or by calling the Milken Exchange at 310-998-2825.

Student Achievement

What little systematic research has been done suggests that classroom technology can raise student achievement and even improve the overall learning environment in schools – but only when it is placed in the right hands and used in the right ways. In fact, when used for the wrong purposes, computers can do more harm than good.

A newly published study by the Educational Testing Service (ETS) in Princeton, New Jersey, found that when used selectively by trained teachers, computers significantly enhanced the performance of middle school students in mathematics. But the study also found the value of computers in elementary school is far more limited and that, when used primarily for drills and practice at either level, computers can be counterproductive.

1 - 88

The ETS report, published in *Education Week's* "Technology Counts '98" issue, is the first large-scale examination of the link between computer use and student achievement. The study draws on a national database of student test scores, classroom computer use and other information, including school climate. Among some of ETS' findings are the following:

- Eighth-grade students whose teachers used computers for "simulations and applications," rather than for "drill and practice," outperformed their peers.
- Students whose teachers had professional development in computers outperformed – by more than one-third of a grade level – students whose teachers did not.
- Students who spent more time on computers in school did not score higher than their peers; in fact, they performed slightly worse.

"What matters most," the ETS study concluded, "are not the machines and the wiring, but what teachers and students do with them."

Education Week's "Technology Counts '98" is available online at www.edweek.org. Or, order a copy for $6 by calling 800-346-1834. The 114-page report includes the most recent national and state-by-state data on technology access, capacity and use.

The Educational Testing Service's policy report, *Does It Compute? The Relationship Between Educational Technology and Student Achievement in Mathematics*, is available online at www.ets.org/research/pic.

Teacher Training and Technical Support

School systems seem to be paying greater attention to the important role played by teacher training. Survey results from the latest National Assessment of Educational Progress (NAEP) show that 81% of the nation's 4th graders had teachers who had received professional development with computers within the past five years. Among 8th graders, 76% had mathematics teachers who had received such training within the past five years.

But while the vast majority of teachers have had some training in education technology, far fewer have had training in more sophisticated uses of technology. Forty percent of teachers, for example, report having had no formal training in using the Internet. And fewer than four in 10 schools report having either a full- or part-time technology coordinator to make sure computers are maintained and updated, purchase new software and provide technical support for teachers.

Thirty-eight states have technology requirements for teacher-preparation programs, but the requirements vary widely in rigor. In Massachusetts, for example, schools of education are required only to help prospective teachers learn to use "new technologies." In Idaho, education schools must assess whether teaching candidates are proficient in technology.

The National Council for Accreditation of Teacher Education's report, *Technology and the New Professional Teacher*, is available by calling 202-466-7496. Also available online at www.ncate.org.

1-89

SCHOOL CHOICE

A RESEARCH EXPERIMENT WITH AT-RISK STUDENTS

Research Summary

School choice deserves closer scrutiny. Kansas, like most states, needs to take a step back and re-think K-12 education. Too many children are not reaching their potential, and the global future promises that they will need a full complement of intellectual skills even more than their parents did. With this monumental task facing us, it is important to ensure the state does not become so mired in its traditional ways of providing education that it misses opportunities for improvement. We must not shy away from implementing and evaluating new approaches in order to discover if they are new routes for student success.

No one knows definitively whether school choice programs benefit children. Even the much-discussed Milwaukee school choice program needs better data to find out whether school choice makes a difference in low-income children's education. It's time to take a closer look. Many have advocated the benefits for school choice and recent articles in magazines like *The Atlantic Monthly* argue that school choice programs have a plethora of benefits for educating our youth with which the spectrum of political factions should agree.

Glasscock-Tanner "Years of Promise -- Kansas' Keys to Success" Education Plan

- The Kansas Department of Education will contract with a professional researcher to design and conduct a 4-year, research-based school choice experiment for at-risk Kansas children in grades 3 through 6. The KSDE will report the findings to the legislature.

- The experiment will be structured to answer the following question: Is there a positive, negative or neutral correlation between vouchers and student learning? The project must have a control group as a benchmark to measure opportunity scholarship students' success in public and private schools versus the success of public school students in their regular district school.

- Opportunity scholarship students and control group students will be tested on national norms tests not currently being used in the Kansas assessments at the beginning and end of the 4-year experiment. The same tests will also be administered each year of the experiment.

- School opportunity scholarships will be equivalent to the base per pupil funding. The weights normally accrued by the public school for the opportunity scholarship student will be used to fund the research on the program. If a private school charges a tuition amount less than the base per pupil funding, the remainder will be placed in the research fund. Opportunity scholarships for special education students will be based on the cost of service as determined by their Individual Education Plan.

- The students involved in the experiment must be representative of the demographics of the state's student population, including special education students. The researcher designing the study will determine the minimum number of students necessary for statistically valid research. Students in the experiment will be selected by lottery.

- To qualify for school choice, students must have been enrolled in public school during the year before their first year of involvement in the program.

SCHOOL CHOICE
RESEARCH HIGHLIGHTS

"After nine years, no one can say with assurance whether or to what extent the [school choice] program has met those objectives. The time has come to make a renewed effort to assess the educational, economic, and social impacts...."

Alex Molnar
"Unfinished Business in Milwaukee"
Education Week

"What if the ability to escape might help to make the schools better? And what if testing this proposition can't make anyone worse off? Yes, big voucher plans may require an act of faith, but it wouldn't be the first gamble in American education to work. A much smaller federal government rolled the dice on land-grant colleges in the 1860s with only a notion of what would happen; the research they sparked made U.S. agriculture the world's most productive."

Matthew Miller
A Bold Experiment to Fix City Schools
The Atlantic Monthly

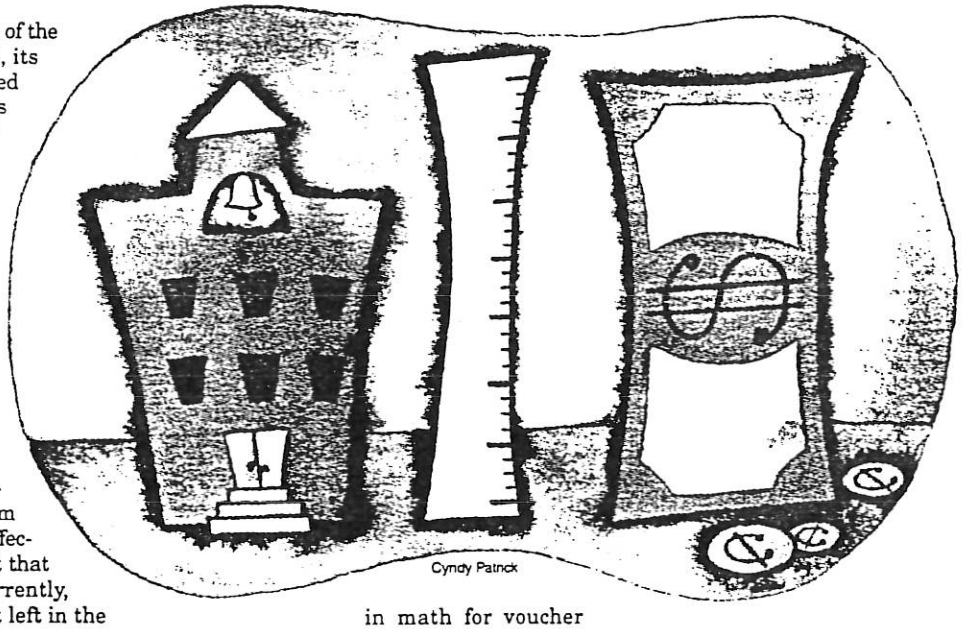
Unfinished Business in Milwaukee

Why Launch a National Voucher Experiment When We Have Not Fully Examined the One We Have

Alex Molnar

In June of 1998, the Wisconsin Supreme Court found constitutional the expansion of the Milwaukee Parental Choice Program to include religious schools. The refusal of the U.S. Supreme Court to review the Wisconsin decision means that, for the time being, the only voucher program in the country that involves religious schools without a constitutional sword over its head is Milwaukee's.

Although the constitutionality of the Milwaukee program is settled, its impact is still a matter of heated debate, a debate that continues to have national significance as state legislatures and the U.S. Congress consider voucher bills. Some disagreements over vouchers are not likely to be resolved by additional data. However, to the extent that data can help shape public policy that frames government-financed private-school-voucher programs, there is much yet to be learned from Milwaukee.



Unfortunately, in 1995, when the Wisconsin legislature expanded the Milwaukee program to include religious schools, it effectively removed the requirement that the program be evaluated. Currently, the only evaluation component left in the law requires the Wisconsin legislative audit bureau to conduct an audit of the program. The audit bureau's report is expected sometime late this year or early next year, but voucher schools are not required to participate in the statewide testing program, to administer the same tests, or to share the testing data they have on their students. Therefore, it is doubtful that the audit bureau's report will have much to say about the relative academic performance of voucher schools.

The legislature's decision to virtually eliminate any evaluation of the performance of the Milwaukee program is puzzling because the effects of the Milwaukee voucher program were (and still are) far from settled. For example, between 1990 and 1995, the University of Wisconsin-Madison political science professor John Witte conducted the annual evaluations required by the original legislation. He found that attending voucher schools conferred no achievement advantage. Using a different approach to the same data, the researchers Jay P. Greene, Paul E. Peterson, and Jiangtao Du concluded that attendance at voucher schools had a cumulative achievement effect that only showed up after three or four years. They reported that students who were in the program three or four years had significantly higher scores in reading and in math than students who applied for but were not admitted to the program. Cecilia Rouse conducted a third analysis of Mr. Witte's data. She found a significant advantage in math for voucher students who were in the program for three or four years and no achievement advantage in reading for voucher students.

Given the differing interpretations of the same data and the small number of students involved, it would have been desirable for the Wisconsin legislature to address the perceived flaws in the data by drafting a tightly drawn and comprehensive evaluation provision when it expanded the program. However, since the legislature did not, both supporters and opponents of educational vouchers can point to the Milwaukee program and claim support for their views. As the program continues and expands, virtually no new data are being systematically collected, and important public-policy questions remain unanswerable.

One of the fundamental arguments of many voucher supporters is that private schools do a better job of educating children than public schools. Roman Catholic schools have, of late, been deemed particularly praiseworthy. It would be helpful to know how Catholic and other private schools participating in the Milwaukee voucher program are

Professor of Ed & Director of Center for Ed. Research, Analysis & Innovation, Univ. of Wisconsin — Milwaukee

1-93

performing, but the Milwaukee Archdiocese refuses to release its test results. Rumor is that its voucher students are performing less well than their public school counterparts. In truth, no one knows. The last time the public saw the archdiocesan test results was in 1991, when *The Milwaukee Journal* succeeded in getting them. Those results suggested that when children of similar social and economic background were compared, Catholic schools were doing no better and perhaps a bit worse at educating minority children than the Milwaukee public schools.

Achievement data are not the only sort of information currently unavailable in Milwaukee. Also missing is information relevant to issues such as racial segregation and social-class stratification. According to the Wisconsin Department of Public Instruction, following the Wisconsin Supreme Court's ruling, the number of private schools participating in the program increased from 23 nonreligious schools in 1997-98 to 86 private schools (30 nonreligious and 56 religious) in 1998-99. The number of students involved increased from 1,500 to 6,050. The program is legally allowed to enroll up to 15,000 students. Given this increase in enrollment (and the potential for future increases), it would be very helpful to know whether Milwaukee's voucher program increases, decreases, or has no impact on school segregation in the city. Right now, no one can say for sure; however, there may be reason to be concerned. Of the 6,050 students the state department of public instruction says were in the program as of January 1999, 2,274 were already enrolled in a private school; 1,114 had never been enrolled in school before; 1,295 had been in the voucher program the preceding year; and 1,367 were enrolled in the Milwaukee public schools the preceding year. No one knows the racial or ethnic characteristics of students who were previously enrolled in private schools, nor those who were never enrolled in school before. It is likely, however, that the majority of students previously enrolled in private schools were white. As time passes and students move into and out of voucher schools, it will be impossible to know what impact the voucher program is having on racial isolation without the ability to gather systematic data.

A good deal of the pro-voucher argument concerns the alleged efficiency of private schools. In Milwaukee, the benefits of the voucher program relative to its cost cannot be clearly determined. It would appear, for example, that since the state now pays up to $4,894 per voucher student, and a majority of voucher students are in the relatively cheap pre-K-through-3 grades, voucher schools may currently enjoy a financial advantage over Milwaukee's public schools. The relative costs of public and voucher schools cannot be accurately calculated, however, without being able to gather and analyze comparable financial data for the schools and school systems involved in the program.

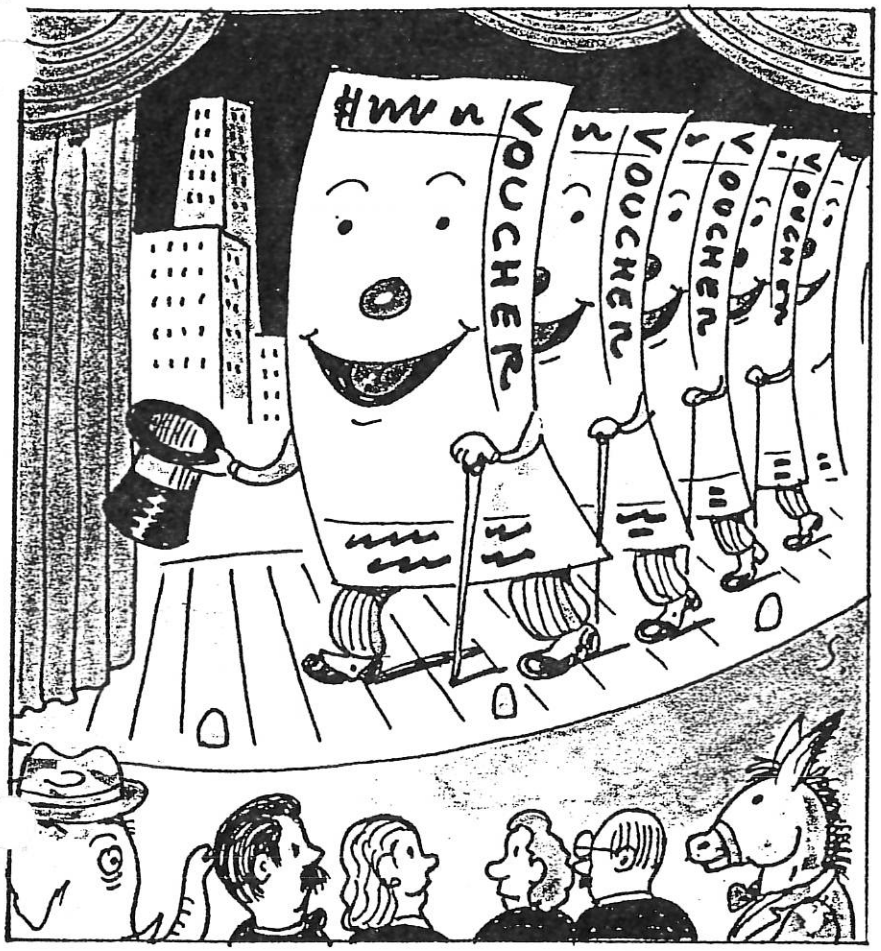
One of the bedrock principles advanced by voucher advocates in Milwaukee and elsewhere is the importance of parents' being able to choose the schools that their children attend. Thus empowered, parents will impose accountability on voucher schools by "voting with their feet." Given the importance of this principle to the voucher reform, it is important to learn the extent to which the Milwaukee voucher program provides genuine as opposed to formal options for parents. There is some evidence to suggest that this is an area that merits careful attention.

This past spring, People for the American Way and the National Association for the Advancement of Colored People sponsored an investigation into the admissions procedures and other practices of Milwaukee's voucher schools. The investigation, conducted by the Metropolitan Milwaukee Fair Housing Council, found instances of unlawful admissions requirements, such as requiring parental agreement to support and engage in religious activities, to be active in school fund raising, and to provide volunteer services. The legal requirement to randomly admit students was not always followed, nor was the right to opt out of religious activity. A number of schools appear to have imposed fees that are illegal under the Milwaukee parental-choice law. Further, some schools have enrollment periods that may operate in such a way as to discourage the enrollment of new voucher students. At the moment, there is no way to know the degree to which such school practices or other factors restrict parent choices.

According to a recent report by the Milwaukee-area voucher advocate Susan Mitchell, when the Milwaukee program was enacted in 1990, it was intended to accomplish three purposes: "to foster more educational options for poor parents, better achievement for their children, and improved performance in the Milwaukee public schools." After nine years, no one can say with assurance whether or to what extent the program has met those objectives. The time has come to make a renewed effort to assess the educational, economic, and social impacts of Milwaukee's voucher program.

Some researchers, policymakers, and interest groups are currently promoting the idea of a national voucher experiment as a way of addressing unresolved questions about educational vouchers. It would be far better to focus on Milwaukee, where a large voucher program already exists and is under no constitutional threat. The time is right for a team of social scientists to approach the Wisconsin legislature with a well-thought-out evaluation plan recognized as fair by both voucher advocates and voucher skeptics. Good public policy demands that the legislature be willing to modify the Milwaukee Parental Choice Program legislation so that such an evaluation could be carried out if such a plan comes forward. ■

1-94



A Bold Experiment to Fix City Schools

A proposal for school vouchers on which Milton Friedman, Lamar Alexander, and Kweisi Mfume, the president of the NAACP, all agree

by Matthew Miller

WHEN Maria Neri's daughter Tina finished eighth grade, two years ago, her scholarship at a Catholic elementary school in south-central Los Angeles ended. The parochial high school in which Neri (not her real name) hoped to enroll Tina charges $3,500 a year—a third less than the $5,400 Los Angeles would spend to educate Tina in public school. Neri, thirty-three, earns $600 a month as a part-time teacher's aide; she's looking for a second, and perhaps a third, job. Her husband, from whom she is separated, earns $1,200 a month as a laborer in a glass factory. He pays his wife's monthly rent of $340, but offers no support beyond that. After paying for food, a phone, gas, and other expenses, Neri had no money left to put toward private school for Tina. Yet she was afraid to send Tina to the neighborhood public school, where the walls were covered with graffiti, and *"cholos,"* or gang members, had been involved in shootings that brought police helicopters to the campus. So Neri used her sister's address to enroll Tina at another public school, which, though twenty minutes away, at least seemed safer. But it is far from ideal. Classrooms each have forty to forty-five children belonging to several different grades. Tina, sixteen, says the teachers often have the students watch movies. Her math teacher was so confused about who Tina was that he gave her an F for not completing many assignments—a grade he changed, with embarrassment and an apology, after Neri confronted him with Tina's completed workbook. "I can see the difference," Neri says. "She's going down." Tina says she would go back to Catholic school if they could afford it. "I talk to my daughter," Neri explains, "and say, 'I'm sorry.'"

Neri's desire to send Tina to a better school is at the heart of one of the nation's most important and most demagogic debates. Through vouchers, often touted as an answer to Neri's problem, the government would give parents some or all of the money it now spends educating their children to use at a school of their choice. Depending on whom you listen to, vouchers are either a lifeline or a death knell. "It is quite simply an issue of survival for our nation's poorest students," says Dan Coats, a Republican and a former senator from Indiana. But Kweisi Mfume, the president of the National Association for the Advancement of Colored People, calls vouchers a "terrible threat," and Sandra Feldman, the president of the American Federation of Teachers, says they mean "a radical abandonment of public schools and public education."

These are heated claims, especially given the relatively small number of students who are involved in voucher programs today. Just over 52 million students attend grades K through 12 in the

Illustrations by J. C. Suarés

1-95

United States. Only two cities offer publicly funded vouchers: in Milwaukee (whose breakthrough program was begun in 1990) roughly 6,000 of 107,000 students get vouchers; in Cleveland about 4,000 of 77,000 do. In May, Florida approved a plan under which students at the poorest-performing schools would get vouchers. Four schools are expected to be eligible this year, and 12,000 of the state's 2.3 million K–12 kids are expected to use vouchers over the next four years. Privately funded voucher programs in thirty-one cities served roughly 12,000 children last year; ten new such programs came into being for the 1998–1999 school year. Two wealthy investors, Ted Forstmann and John Walton, recently announced a plan to fund (along with other donors) $170 million in vouchers, which will reach 40,000 new students over the next four years.

Add these numbers up and you get 74,000 children—about 0.1 percent of students. Add 200,000 for those students in the 1,200 charter schools around the country (which also give parents a choice), and the proportion comes to only 0.5 percent of schoolchildren. In other words, the school-choice debate is taking place utterly at the margins. At this rate, for all the fuss, it's hard to imagine that any impact could be made on the skills and life chances of students stuck in our worst public schools in time to prevent what the Reverend Floyd Flake, a voucher advocate and a former Democratic congressman from New York, calls "educational genocide."

Where is the "voucher left"? Vouchers have a long but unappreciated pedigree among progressive reformers.

This tragedy is most pronounced in big cities, whose public schools together serve six million children. Despite heroic local efforts and pockets of success, depressing evidence mounts of an achievement gap between students in cities and those in suburbs, where, school-watchers say, most schools are doing fine, largely because they're safer, better funded, and less prone to the social ills that plague cities. Of Detroit's eleventh-graders 8.5 percent were deemed "proficient" in science on Michigan's 1997 statewide exam. Fourth-graders in Hartford were a tenth as likely as Connecticut students overall to show proficiency on the state's three achievement tests in 1996. Only two percent of Cleveland's minority tenth-graders have taken algebra. "The numbers tell a sad and alarming story," a special report in *Education Week* concluded last January. "Most 4th graders who live in U.S. cities can't read and understand a simple children's book, and most 8th graders can't use arithmetic to solve a practical problem." As polls prove, increasing numbers of urban parents like Maria Neri want a way out. It seems immoral to argue that they must wait for the day when urban public schools are somehow "fixed." It's even harder to argue that bigger voucher programs could make things worse.

Yet a political standoff has kept vouchers unavailable to nearly 99 percent of urban schoolchildren. Bill Clinton and most leading Democrats oppose them, saying we should fix existing public schools, not drain money from the system. Teachers' unions, the staunchest foes of vouchers, are among the party's biggest donors, and sent more delegates to the 1996 Democratic National Convention than did the state of California. Republicans endorse vouchers as a market-based way to shake up calcified bureaucracies, but they generally push plans that affect only a few students. The distrust that has led to today's gridlock is profound. Republicans view Democrats as union pawns defending a failed status quo; Democrats think Republicans want to use urban woes as justification for scrapping public education and the taxes that fund it.

MISSING entirely from the debate is the progressive pro-voucher perspective. To listen to the unions and the NAACP, one would think that vouchers were the evil brainchild of the economist Milton Friedman and his conservative devotees, lately joined by a handful of desperate but misguided urban blacks. In fact vouchers have a long but unappreciated intellectual pedigree among refo ers who have sought to help poor chi and to equalize funding in rich and p districts. This "voucher left" has always had less cash and political power than its conservative counterpart or its union foes. It has been ignored by the press and trounced in internecine wars. But if urban children are to have any hope, the voucher left's best days must lie ahead.

Finding a productive compromise means recalling the role of progressives in the history of the voucher movement and exposing the political charades that poison debate. It means finding a way for unorthodox new leaders to build a coalition—of liberals for whom the moral urgency of helping city children trumps ancient union ties, and of conservatives who reject a laissez-faire approach to life's unfairness. The goal of such a coalition should be a "grand bargain" for urban schools: a major multi-year test of vouchers that touches not 5,000 but 500,000 children, and eventually five million—and *increases* school spending in process. The conventional wisdom that today's whittled-down pilot programs are all that is politically achievable. The paradox is that only through bigger thinking about how vouchers might help can a durable coalition emerge.

IN 1962 John E. "Jack" Coons, an idealistic thirty-two-year-old law professor at Northwestern University, was asked by the U.S. Civil Rights Commission to study whether Chicago schools were complying with desegregation orders. Coons soon found that what really interested him was a different question: Why were suburban schools so much better than those downtown? Over the next few years Coons, eventually joined by two law students, Stephen Sugarman and William Clune, found one answer in what would become a source of enduring outrage: America's property-tax-based system of public-school finance created dramatic disparities in the resources available to educate children.

This financial aspect of education's vaunted tradition of "local control" rarely the subject of national controversy. In part that is because it gives the nation's most powerful citizens both lower taxes *and* better schools. Imagine two towns, Slumville and Suburbia. Slumville has

$100,000 in taxable property per pupil; Suburbia has $300,000. If Slumville votes to tax its property at four percent, it raises $4,000 per pupil. But Suburbia can tax itself at two percent and raise $6,000 per pupil. Suburbia's tax rate is half as high, but its public schools enjoy 50 percent more resources per student.

In the 1960s affluent districts routinely spent twice what nearby poorer ones did, and sometimes four or five times as much. To Coons and his colleagues, such inequity in a public service was indefensible. Beginning with *Private Wealth and Public Education*, a book that he, Sugarman, and Clune published in 1970, Coons has denounced the system eloquently. It's worth sampling his arguments, because the left's case for choice is usually drowned out by the right's cheerleading for markets, or by urban blacks' cry for help. In a 1992 essay, "School Choice as Simple Justice," Coons wrote,

> This socialism for the rich we blithely call "public," though no other public service entails such financial exclusivity. Whether the library, the swimming pool, the highway or the hospital—if it is "public," it is accessible. But admission to the government school comes only with the price of the house. If the school is in Beverly Hills or Scarsdale, the poor need not apply.

Coons's point was simple: the quality of public education should not depend on local wealth—unless it is the wealth of a state as a whole. "Everyone ought to be put in a roughly equivalent position with regard to what the state will do," Coons, now an emeritus law professor at Berkeley, says.

Coons and Sugarman made a successful case for the unconstitutionality of the school-finance system in California's famous *Serrano* case in 1971, beginning a national movement to litigate for school equity. Although it was little noticed then, they cited vouchers as a potential remedy. The idea was to give courts a way to instruct legislatures to fix things without having to mess with local control. Asking legislatures to centralize school funding at the state level was a political nonstarter. But through various formulas, Coons and Sugarman argued, the state could give families in poorer districts enough cash in the form of vouchers to bring education spending in those districts up to that of better-off districts. And what could be more "local," they reasoned, than giving families direct control over the cash to use at schools as they chose?

COONS and Sugarman, focusing on school equity, thus arrived at a policy that Milton Friedman had been urging through a principled commitment to liberty and to its embodiment, the market. Friedman's 1955 essay "The Role of Government in Education" is viewed as the fountainhead of the voucher movement. In an ideal world, the future Nobel laureate reasoned, the government might have no role in schooling at all; yet a minimum required level of education and its financing by the state could be justified.

> A stable and democratic society is impossible without widespread acceptance of some common set of values and without a minimum degree of literacy and knowledge on the part of most citizens . . . the gain from the education of a child accrues not only to the child or to his parents but to other members of the society. . . . Yet it is not feasible to identify the particular individuals (or families) benefitted or the money value of the benefit and so to charge for the services rendered.

However, Friedman said, if this "neighborhood effect" meant that the government was warranted in paying for K–12 education, another question remained: Should the government run the schools as well? Friedman's view was that schools could be just as "public" if the government financed but didn't administer them. That notion remains virtually unintelligible to leaders in public education, perhaps because it is so threatening.

Friedman's analogy (adopted by every voucher proponent since) was to the G.I. Bill, which gave veterans a maximum sum per year to spend at the institution of their choice, provided that it met certain minimum standards. Likewise, for elementary and secondary schooling Friedman envisioned a universal voucher scheme that would give parents a fixed sum per child, redeemable at an "approved" school of their choice. Such a school might be nonprofit or for profit, religious or secular. Parents could add to the sum if they wished. The role of government would be limited to assuring that "approved" schools included some common content in their programs, "much as it now inspects restaurants to assure that they maintain minimum sanitary standards." In Friedman's view, market-style competition for students would spur the development of schools that were better tailored to families' needs and cost less than those run by notoriously inefficient public bureaucracies.

Friedman's and Coons's different angles of vision represent the ancient tug between liberty and equality within the pro-voucher camp—a debate the two have waged since Friedman was an occasional guest on Coons's Chicago radio show, *Problems of the City*, in the 1960s. Friedman today isn't bothered by issues of school-finance equity. "What's your view of inequity in clothing and food?" he snapped when asked recently, saying that such concerns reflect Coons's "socialistic approach." And even if public schools were making every child an Einstein, Friedman says, he would still want vouchers. "Private enterprise as opposed to collectivism," he says, "would always be better."

Coons is less ideological. In his view, choice would improve the public schools, which he believes would always be chosen by the majority, even with a full-blown voucher system. The prospect of losing students (and thus funding) would force improvements faster than today's seemingly endless rounds of ineffectual education fads. If poor children got a decent education under the current system, he adds, he probably wouldn't have devoted his life to these issues.

The fate of disadvantaged children under a voucher regime is where the Coons-Friedman clash is sharpest. Coons would be glad to offer vouchers to all low-income students and to no one else if such a step were necessary for consensus. He fears that under a universal voucher system they could get left behind, as schools competed to recruit better-off, smarter, healthier (nondisabled) students. The incentives are plain: such children would be easier to teach, and schools could charge wealthy families far more than the voucher amount to maximize profit. Coons and the voucher left therefore insist that any universal scheme should include protections for low-income and disabled children. Examples would be increasing the voucher amount for those children to make them more attractive to schools, and letting schools redeem their vouchers only if, say, 15

(Continued on page 26)

1-97

(Continued from page 18)
percent of new places were reserved for such children, for whom the voucher would cover tuition. To Friedman, these are unacceptable intrusions on schools' freedom to operate as they like, turning vouchers into "a welfare program, not an education program."

WITHOUT a link to unions—which, despite the waning of their influence, remain one of the few sources of progressive ideas in American public life—liberal pro-voucher champions have had little political impact. The muting of their voice, combined with the ease of legislating pilot programs, explains why few urban children have a choice today. What's more, deceptive arguments by both teachers' unions and conservative activists keep the broader public confused.

A "grand bargain": combine a bigger road test for vouchers with increased per-pupil spending.

Teachers' unions (and voucher foes generally) rely on five dubious arguments.

There's no evidence that vouchers work. The trials have been so isolated, unions say, that their results are unproved. That's a nervy case to make when it is union opposition that has kept the trials small. Pro- and anti-voucher forces have funded research in Milwaukee and Cleveland that purports to show why Johnny is doing demonstrably better or worse under vouchers. It is impossible to make sense of these dueling studies, whose sample sizes are so small that results seem to turn on whether, say, three children in Cleveland handed in their homework on time. Wealthy conservatives are now offering vouchers to all 14,000 at-risk children in a poor San Antonio district in part so as to compile a broader database from which to judge the impact of voucher systems. (In the first semester of the program 566 children taking vouchers left district schools.) For now the "no evidence" argument says more about union chutzpah than about voucher performance.

Vouchers drain money from public schools. Sandra Feldman, of the American Federation of Teachers, says that the $10 million Cleveland uses to give vouchers to 4,000 children would be better spent on measures that would benefit *every* child, such as shrinking class sizes and launching proven reading programs. But this is disingenuous. Cleveland provided the $10 million in addition to more than $600 million in existing school spending in order to mollify unions, which insisted that vouchers not "come out of the hide" of public schools. It's unfair for unions to turn around and complain that the extra cash they insisted on should have gone elsewhere. The truth is that public schools are free to fund such measures now by shifting priorities within their budgets. And when broader voucher plans let the amount that public schools receive per student follow students who leave the system, the public-school coffers are not drained—schools receive the resources their enrollment merits.

Vouchers are unconstitutional. Some critics say that voucher use at religious schools violates the Constitution's ban on "establishment of religion," but the better view of the Supreme Court's confusing jurisprudence here suggests that's wrong. After all, no one thinks that federal student loans are unconstitutional when they are used by students to attend Notre Dame. Last June, Wisconsin's highest court upheld Milwaukee's plan, because the voucher goes to parents to use where they like, not to any particular type of school. In union hands, moreover, this legal complaint seems suspiciously tactical. It can't be that we are constitutionally obligated to imprison urban children in failing schools.

The capacity isn't there. Public schools serve 46 million K–12 children, private schools six million. Since private schools can't accommodate more than a fraction of today's students, opponents say, vouchers can't be a meaningful part of school reform. "Where are these schools going to come from?" Sandra Feldman repeatedly asked during an interview with me.

The first response to this argument is to ask, Then what's the problem? If as a practical matter unions feel that mo children with vouchers will remain wh they are, it's hard to see what the harm in trying them. A second response is that even relatively few defections from public schools may spur efforts to improve them. Districts with innovative charter schools have reported such a reaction.

The larger answer, however, is that broader voucher schemes would prompt many institutions and entrepreneurs to add schools and spaces to the "market." This would happen not overnight but over a number of years. The initial spaces would be likely to come from Catholic schools, which account for half the private-school slots in the country. Jerome Porath, the schools chief for the Los Angeles archdiocese, says that if every student got a voucher worth an amount close to the current per-pupil expenditure in California, over several years enough facilities could be built or rented "to accommodate everybody who wanted to come." "We'll get out our spreadsheets and figure it out," he says. Milton Friedman adds, "You can't think of it in terms of the existing stock of schools. There will be a flood of new schools started."

Profit is bad. Voucher foes act as if there were something venal about the profit motive when applied to schools. But public education is already big business. The $320 billion spent last year on K–12 schooling is lusted after by textbook publishers, test designers, building contractors, food and janitorial services, and software companies, to name only a few examples. This largesse inevitably brings scandals—for example, the California flap in 1996 over whether campaign contributions influenced a big textbook purchase. Like health care, defense, and other major public services, schools will always be partly about business; vouchers would simply change who controls the flow of cash. There's no reason to think that the abuses under a voucher system would be worse than abuses today.

Voucher foes make other unpersuasive claims. They say that vouchers will cream off the most-talented children and the most-active parents—a worry that seems acute primarily because today's voucher plans remain tiny. They say that private schools will unfairly be able to avoid troublemaking kids by not admitting them—ignoring the fact that public

1-98

districts themselves often send such kids to special schools of "last resort." They say the oversight that will follow public money will make private schools resemble public bureaucracies—ignoring the greater flexibility that most analysts say such schools will retain in hiring and firing, resource allocation, and curriculum design. Finally, they argue that it is crazy to subsidize more-affluent parents who already pay for private school—a seemingly powerful charge until one recalls that such families are now paying twice for schools, and that vouchers offered only to poor families would avoid the problem entirely.

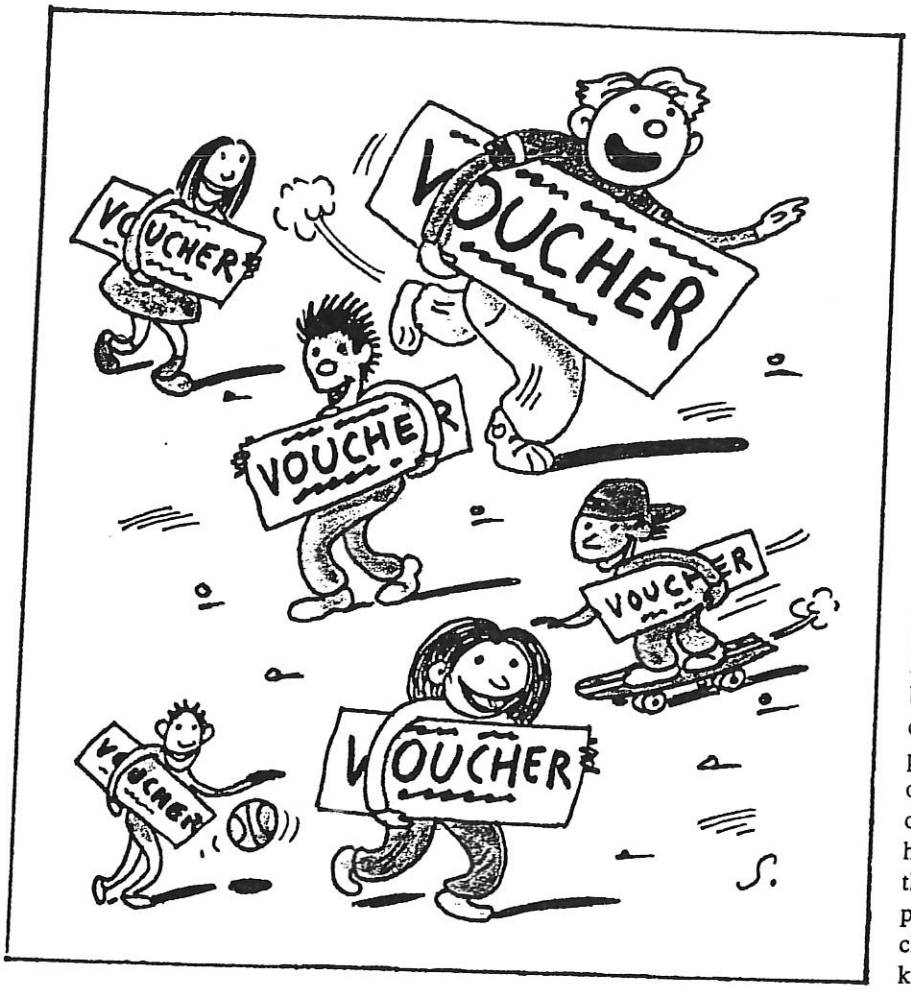
For their part, conservative voucher fans peddle one big misconception: vouchers can save lots of money because per-pupil spending in private schools is typically less than half that in public schools today. It is true that religious schools have fewer administrators and lower-paid teachers, and invest less in such amenities as theaters, labs, and gymnasiums. But private schools don't have to take costly disabled and "special education" children; and often public schools must offer extras such as English as a Second Language, breakfast and lunch programs, and transportation. When such differences are taken into account, and hidden subsidies for church space and staff in religious schools are counted, the gap shrinks. Coons says that a voucher's value needs to be no lower than 85 percent of total per-pupil spending in order to stimulate capital investment in new schools. Set it too low, and the result will be simply to fill the handful of empty Catholic-school seats.

The right's claim that vouchers will deliver big savings also ignores the case for spending more in many big cities, where dilapidated buildings may collectively require as much as $50 billion in repairs. Some public school bureaucracies—Washington, D.C., and St. Louis come to mind—seem so hopeless that it would be senseless to pour new money in until management has improved. But despite run-down buildings and higher proportions of special-needs students, cities such as Philadelphia and Baltimore spend substantially less per pupil than do their states overall.

DISINGENUOUS rhetoric, visceral distrust, maximal posturing, minimal progress. Political debates escape this kind of dead end when grassroots pressure makes the status quo untenable, or when leaders emerge with fresh ways of framing the issues. It's possible that urban schools will fall so far that the poor revolt; or crime, bred by ignorance, might worsen in ways that force society to act. There's a better path to hope for, however, if new leaders can teach us to think differently about today's predicament.

Sounds of rethinking and compromise are in the air. Arthur Levine, the president of Columbia University's Teachers College, is a lifelong liberal and a voucher foe. Yet, frustrated by the seemingly hopeless troubles of inner cities, Levine called last June for a "rescue operation" that would give vouchers to two to three million poor children at the worst urban public schools. "For me," Levine says, "it's the equivalent of Schindler's list." Lisa Graham Keegan, Arizona's superintendent of public instruction and a rising Republican star, calls the property-tax base for school fi-

1-99



nance "pernicious" and "wholly unfair." She wants a system of "student-centered funding," in which revenues from a source other than property taxes would be distributed by the state on an equal per-pupil basis through vouchers.

If leading liberals are willing to question the public school monopoly, and prominent conservatives hear the call of justice, the voucher debate has a chance to move forward. The sensible first step would be a much bigger road test. Here's the idea I have put to various players in the debate: Suppose everyone came together and said, Let's take three or four big cities where we agree the public schools are failing. (Leave out dens of mismanagement like Newark and Washington, where spending is high but ineffective.) In these cities we'll raise per-pupil spending by 20 percent, giving urban schools the resources the left says they need, and thus going far to achieve the Coons vision of funding equity. But we'll implement this increase by way of a universal voucher system that finally gives *every* child a choice. So, for example, in a city that now spends $5,000 per pupil, every child would get a $6,000 voucher.

Such a proposal, serving half a million children, would cost $660 million a year. If the voucher system were then extended to all six million big-city children (a logical step if results of the trial were promising), the price tag would be $8 billion a year, or 0.4 percent of federal spending. (For purposes of discussion, I left aside the question of who outside the district would fund the 20 percent increase, though the surplus-rich federal government comes readily to mind.) The responses to this idea suggest how quickly the scale of today's debate could change—and who is responsible if it doesn't.

Jack Coons, the "egalitarian," said it sounds great. Clint Bolick, a conservative lawyer who is active in the voucher movement, also thought it could work—though, he said, the spending increase would mean that "some of my fellow conservatives would have apoplexy." Polly Williams, who led the drive to enact vouchers in Milwaukee, was anxious about extending them to students who aren't poor, so we agreed to give them only to children eligible for the federal school-lunch program. This would still get vouchers to 78,000 children in Milwaukee instead of the current 6,000, and to four million city children nationwide. We would move pretty far toward universal coverage this way, since, sadly, two out of three city children qualify for school-lunch assistance.

What about the NAACP? To date the organization has welcomed philanthropic efforts, but when public funds are at issue, it stands by the unions. Julian Bond, the chairman of the NAACP, recently called vouchers "pork for private schools." Yet when I asked Kweisi Mfume, the NAACP president, about this proposal, he didn't hesitate. "I don't have a problem with that at all," he said. Mfume says that NAACP opposition has been not ideological but based on three concerns: the association doesn't want programs that leave nearly every child out; it wants accountability to the public on student performance; and it wants an honest approach to higher costs—such as those for transportation—that must be paid to make the system work for poor children. The pilot programs in Milwaukee and Cleveland fail especially on grounds one and three; the bargain I sketched addresses them. Mfume said he was open to the proposal as long as the NAACP's concerns were met, even if that meant taking a stance different from the unions'.

"It's a bad idea," Milton Friedman said at first, arguing that any increase in spending would "fuel the racketeers in the education business." Friedman's point is that raising spending could create further opportunities for profit-hungry operators to take the vouchers and run schools much more efficiently—not to their benefit. Owing to systematic federal overpayments, Medicare HMOs face just such scams in many places today.

But outliers like Washington, D.C., aside, it's not clear that urban schools are overspending. Given that, isn't it worth running a little risk to get a substantial voucher test under way? It seemed that Friedman wouldn't sign on, but toward the end of our discussion he relented. "I'll tell you what I would go for," he said. Friedman has always believed that so many families would flee public schools if given a voucher worth even half what is now spent per pupil that resources for

each child remaining in the system would rise. (If ten public school children have $5,000 spent on each of them, and three leave taking $2,500 each, spending on the seven remaining would rise about 20 percent, to just under $6,100.) So he would approve of a 20 percent increase in per-pupil spending for those who remained, as long as the voucher was worth only half that. Since Friedman thinks that this 20 percent increase will come over time anyway, he's not compromising his ideals. His principled accommodation is to put his money where his beliefs are and increase spending up front as part of the deal.

But look where we are. Baltimore spends $6,400 per pupil today—versus $6,800 spent by Maryland overall. According to Mfume's reasoning, the NAACP would accept a citywide voucher at roughly $7,600. Friedman could live with $7,600 for current public school pupils but would want a voucher for departing students at $3,800. Surely there's a deal to be made here—and a chance, therefore, to help millions of children while meaningfully evaluating voucher efficacy, addressing questions about everything from student achievement to private profiteering.

What about the politicians? Lamar Alexander seems the likeliest to raise these issues thoughtfully in the 2000 election campaign; as a former Tennessee governor and the Secretary of Education under George Bush, he knows more about schooling than any other presidential aspirant. He has also been down this road before. Alexander bears scars from his ill-fated 1992 struggle to enact a voucher test at the federal level. Called the G.I. Bill for Kids, the plan would have spent $500 million in new federal dollars to give the parents of half a million low- and middle-income children each a $1,000 voucher to use at the schools of their choice. Alexander wagered (correctly) that conservative groups would be content with tiny sums of new money to get their foot in the door, and (incorrectly) that new cash for schools would be something the unions couldn't be seen opposing. In a Democratic Congress the bill went nowhere. Today Alexander says he would urge states to shift toward child-centered funding. And he'd go to Congress with an updated version of Bush's 1992 bill, featuring $1,500 per voucher and an overall $1 billion price tag.

I asked Alexander if he wasn't thinking too small: $1,500 vouchers would be nowhere near sufficient to spark the creation of new schools. And with vouchers spread thin across the country, he would get no trial of how broad-based choice can improve schooling in a community. Why not try the 20 percent spending boost in exchange for universal vouchers in a few cities?

The voluble Alexander went silent for perhaps fifteen seconds as he considered whether to go on record in favor of a policy that would raise spending substantially—something that conservative primary voters would reject.

At length he said yes. Higher per-pupil spending wouldn't be his preferred solution, of course, but if that's what it took to get a bold voucher plan into failing cities, he'd live with it. "I would go high because the stakes are high," he explained, "and to expose the hypocrisy of the unions. If I told the National Education Association that we'd double it in the five largest cities, they wouldn't take it."

Was he right? I met with Bob Chase, the president of the National Education Association, in the union's headquarters in Washington. He made the familiar case

Democrats should see large-scale urban voucher programs as an opportunity, not a threat.

for why vouchers are ineffectual today and would be a threatening distraction for public schools if tried more broadly. Only 25 percent of the adult population has children in the schools, he explained. We need to help the other 75 percent understand why financial support of schools is important. In this regard I sketched the deal: a handful of cities, higher spending, but only through vouchers. My tape recorder captured the staccato response.

"Is there any circumstance under which that would be something that . . ."

"No."

". . . you guys could live with? Why?"

"No."

"Double school spending . . ."

"No."

". . . in inner cities?"

"No."

"Triple it . . ."

"No."

". . . but give them a voucher?"

"'Cause, one, that's not going to happen. I'm not going to answer a hypothetical [question] when nothing like that is ever possible."

"But teachers use hypotheticals every day."

"Not in arguments like this we don't. . . . It's pure and simply not going to happen. I'm not even going to use the intellectual processes to see if in fact that could work or not work, because it's not going to happen. That's a fact."

Sandra Feldman was similarly unwilling to consider such a plan. If new money is available for cities, both said, it should be spent to improve the existing system. They would fund pay raises to attract teachers to work downtown, turnaround programs for troubled schools, and general urban programs for health, nutrition, and parenting skills. Of course, pay raises—or smaller class sizes, or any specific reform—could happen under vouchers, if that's what schools felt was needed to attract students.

If one believes that urban education won't improve under the same approach that has failed for years, the path to progress through vouchers follows a simple logic. A progressive hand is needed to pursue the benefits of vouchers without risk to the poor. A number of conservatives are open to such efforts if they make possible larger voucher trials. Given the disastrous state of many urban schools, the Democratic Party should be the natural home of this progressive influence. It is not, because teachers' unions loom large in Democratic fundraising and campaigns. Yet the Republicans' commitment to minorities will probably never be trusted to carry this issue alone. And, not unreasonably, Republicans are unlikely to increase spending for urban schools without ensuring that such increases are tied to systemwide reform.

Changing the Democratic Party's approach to vouchers is therefore the only way to do something serious for urban children anytime soon. This conclusion begets another political syllogism, and an opportunity. Most observers believe that if the NAACP embraced vouchers, it would force the unions to reassess their opposition. Teacher intransigence is sustainable only as long as minority leaders support it, because the children whose future is being blighted are mostly black and Hispanic. Yet as Kweisi Mfume makes clear, getting the NAACP to change its stance would require voucher plans much bolder and more comprehensive than today's pilots.

Thus thinking bigger makes progress likelier. "That's why I've taken the more radical side," explains Floyd Flake, who quit Congress to run his church school and pursue these issues. "It's the only way to force the debate."

At some level even the unions know that their stonewalling is indefensible. "I would never argue with an individual parent who wanted to figure out a way to get his or her child into a better situation," Sandra Feldman says. "But to me, as a matter of public policy, that's not a good argument. The objective is to make the schools good—not to escape them."

But what if the ability to escape might help to make the schools better? And what if testing this proposition can't make anyone worse off? Yes, big voucher plans may require an act of faith, but it wouldn't be the first gamble in American education to work. A much smaller federal government rolled the dice on land-grant colleges in the 1860s with only a notion of what would happen; the research they sparked made U.S. agriculture the world's most productive. The G.I. Bill helped to spawn the postwar middle class. The moral urgency of today's voucher gamble is much greater. For all these reasons, Democrats should see large-scale urban voucher programs as an opportunity, not a threat. After all, once they embraced such a grand bargain, Democrats would be in the driver's seat. They retain, at least for now, the moral authority to speak in behalf of the disadvantaged, and Republicans would not be able to shrink from solutions they have long sought. The alternative is a Democratic Party that favors its funders at the expense of its constituents.

1-102

YEARS OF PROMISE

Kansas' Keys To Success

Glasscock-Tanner Education Plan

Fiscal Note

FISCAL NOTE in millions

	FY 2001	FY 2002	FY 2003	FY 2004
Basic Skills Standards	$0.100	–	–	–
School Performance Incentives	–	$0.650	$0.650	$0.650
*4-year-old at-risk program**	$5.800	$6.800	$7.800	$7.800
*All-day Kindergarten**	–	$15.200	$17.000	$17.900
Teacher Mentor Program	$1.000	$1.000	$1.000	$1.000
Mentor Training	$0.100	$0.100	$0.100	$0.100
National Board Incentives	$0.030	$0.050	$0.080	$0.110
Teacher Scholarships	–	$0.500	$0.500	$0.500
TOTAL	**$7.030**	**$24.300**	**$27.130**	**$28.060**
Children's Initiative Fund	$5.800	$14.400	$16.300	$16.750
SGF	$1.230	$9.900	$10.830	$11.310
Total	$7.030	$24.300	$27.130	$28.060

**The full cost of the 4-year-old at risk program and half of the cost of the all-day kindergarten program is funded from the Children's Initiative Fund. The Children's Initiative Fund is the fund established by the legislature to spend tobacco settlement proceeds on children.*