DOCUMENT RESUME

ED 364 625 UD 029 603

AUTHOR Darling-Hammond, Linda
TITLE Federal Policy Options for Chapter 1: An Equity Agenda for School Restructuring.
INSTITUTION Columbia Univ., New York, NY. Teachers Coll. National Center for Restructuring Education, Schools and Teaching.
SPONS AGENCY DeWitt Wallace/Reader's Digest Fund, Pleasantville, N.Y. Business Roundtable.
PUB DATE Sep 93
NOTE 32p.
AVAILABLE FROM NCREST, Box 110, Teachers College, Columbia University, New York, NY 10027 ($8 pre-paid).
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Access to Education; Children; *Compensatory Education; Disadvantaged Youth; Economically Disadvantaged; Educational Legislation; Educational Policy; Elementary Secondary Education; *Equal Education; Evaluation Methods; *Federal Programs; Financial Support; Inservice Teacher Education; *Policy Formation; *Program Improvement; Resource Allocation; *School Restructuring; Staff Development
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Equal Access; Policy Issues; Reform Efforts

ABSTRACT

This paper suggests specific changes in Chapter 1 legislation (up for reauthorization in the fall of 1993) to make it more responsive not only to schools and students, but also to current school restructuring initiatives and new understandings of teaching, learning, and assessment. Recommendations are made in the areas of greater equity in access to education, serving the whole child and school, teacher preparation and staff development, new evaluation strategies and standards, how funds are allocated to schools and students, and efforts regarding program improvement. Among the recommendations offered are the following: (1) eligibility rules should be changed to allow more schools access to whole-school projects with necessary safeguards and supports for careful planning and appropriate services; (2) equal services and resources must be assured and provided by participating states and school districts to all schools; (3) significant funding should be provided through Chapter 1, perhaps as a proportion of state and local education agency grants for staff development for teachers and administrators through schoolwide professional development programs aimed at improving the overall caliber of teaching as a preventive to remediation; and (4) funds should be allocated to schools strictly on the basis of their concentrations of children living in poverty, with even more emphasis on targeting funds to schools located in high-poverty areas. (Contains 67 references.) (GLR)

Federal Policy Options for Chapter 1: An Equity Agenda for School Restructuring

Linda Darling-Hammond

NCREST

The National Center for Restructuring Education, Schools, and Teaching (NCREST) was created to document, support, connect, and make lasting the many restructuring efforts going on throughout the nation. NCREST's work builds concrete, detailed knowledge about the intense and difficult efforts undertaken in restructuring schools. This knowledge is used to help others in their attempts at change, to begin to build future education programs for school practitioners, and to promote the policy changes that will nurture and encourage needed structural reforms. The Center brings together many voices: those of practitioners and researchers, parents and students, policy makers and teacher educators.

NCREST is supported by a major grant from the DeWitt Wallace-Reader's Digest Fund. NCREST's work in New York City, through its Center for School Reform, is supported by the Leon Lowenstein Foundation and the Aaron Diamond Foundation. Other funders have included the Center for Collaborative Education, the Danforth Foundation, the Geraldine R. Dodge Foundation, the Ford Foundation, the Fund for New York City Public Education, Impact II, the Lily Endowment, Inc., the Andrew Mellon Foundation, the Metropolitan Life Foundation, the National Center for Research on Vocational Education, the New York Community Trust, the New York State Department of Education, and the Regional Laboratory for Educational Improvement of the Northeast and Islands.

Additional copies of this publication may be ordered for $8 each. All orders must be prepaid by check or money order payable to NCREST. Contact:

NCREST
Box 110, Teachers College
Columbia University
New York, NY 10027
Fax: (212) 678-4170

Federal Policy Options for Chapter 1:

An Equity Agenda for School Restructuring

Federal Policy Options for Chapter 1:

An Equity Agenda for School Restructuring

Linda Darling-Hammond

September 1993

Summary

Chapter 1 of the Education Consolidation and Improvement Act of 1981, originally called Title I of the Elementary and Secondary Education Act of 1965, was intended to provide equity in educational opportunity for all of the nation's children, to level the educational playing field between the poor and the privileged by providing needed programs, resources, and services to low-income and low-achieving students. To some extent, Chapter 1 has succeeded in narrowing the gap in basic skills achievement between the educationally advantaged and disadvantaged. At the same time, however, the law's regulations and auditing requirements have produced more bureaucratization and compartmentalization of programs and greater segregation and fragmentation in the delivery of services, creating new problems and educational inequities for the very schools and students Chapter 1 was designed to help.

How can we make the program more effective? This paper suggests specific changes in Chapter 1 legislation -- up for reauthorization in the fall of 1993 -- to make it more responsive not only to schools and students, but also to current school restructuring initiatives and new understandings of teaching, learning, and assessment.

- **Serving the Whole Child and the Whole School**

  1. Eligibility rules should be changed to allow more schools access to whole-school projects with necessary safeguards and supports for careful planning and appropriate services.

  2. "High impact" schools -- those with many children eligible for various federally-funded services, such as bilingual education and special education, as well as Chapter 1 -- should be enabled and encouraged to combine funding from several federal programs.

  3. Initiatives should seek to enhance the capacity of schools to combine funding for education, health, and social welfare services to provide a workable safety net for children in Chapter 1 and other federal categorical programs.

- **Pursuing Greater Equity in Access to Education**

  4. Equal services and resources must be assured and provided by participating states and school districts to all schools. Among these services, perhaps the most important is highly qualified teachers, not just for Chapter 1 services, but for *all* classrooms.

5. States and school districts should seek to ensure that challenging curriculum opportunities and up-to-date curriculum materials and equipment are equally distributed to *all* schools.

- **Teacher Preparation and Staff Development**

6. Significant funding should be provided through Chapter 1, perhaps as a proportion of state and local education agency grants, for staff development for teachers and administrators through schoolwide professional development programs aimed at improving the overall caliber of teaching as a preventive to remediation.

7. Funding strategies for targeted teacher preparation and recruitment should be developed to ensure the availability of highly qualified teachers to Chapter 1 schools with the highest concentrations of low-income children.

- **New Evaluation Strategies and Standards**

8. To make any inference about program contributions to student learning, individual data for only those students who have been in the program over the time period assessed should be used. Student data should be coupled with information about the nature and appropriateness of instructional practices in the program and the school.

9. Assessments of student learning should be multifaceted, represent real performance tasks, and evaluate student progress on the basis of longitudinal gains. To qualify for Chapter 1 funds, schools and districts should be required to show that they have an assessment *system* for evaluating student progress and program accomplishments that

* Measures progress toward the full range of program goals for students;

* Is as authentic as possible (represents real performance tasks and situations);

* Measures higher-order skills and performance abilities;

* Provides multiple methods and occasions for students to demonstrate their skills and knowledge, including many different kinds of performance measures and assessment strategies;

* Allows teachers to evaluate student growth in a cumulative fashion, over time,using several kinds of evidence (e.g., samples of work, observations, performance on tasks);

* Allows assessment of student progress using long-range gains on measures that indicate not how well students do relative to one another, but their ability to perform specific tasks.

- **How Funds are Allocated to Schools and Students**

10. Funds should be allocated to schools strictly on the basis of their concentrations of children living in poverty, with even more emphasis on targeting funds to schools located in high-poverty areas. When districts allocate funds based on the number of low-scoring students in each school, schools can lose their Chapter 1 money as their students start to improve their performance.

- **Program Improvement**

11. Program improvement funds should be used to enable outside teams of expert reviewers to work with school staff in evaluating their practices, developing a plan, and supporting the staff development needed to implement new, more effective program models.

12. Diagnosis must include assessments of district or state- imposed problems that may be undermining school performance. Remedying such factors should be part of the district and state responsibility in overseeing and supporting the plan's implementation.

Acknowledgements

This paper is part of an effort to rethink Chapter 1, coming up for reauthorization by Congress in the fall of 1993. A shorter version appeared in ***Federal Policy Options for Improving the Education of Low-Income Students***, Volume II, ***Commentaries***, edited by Iris C. Rotberg and James C. Harvey, with Kelly Warner, published by The RAND Corporation, Santa Monica, CA. (MR-210-LE)

Special thanks to NCREST communications director Diane Harrington and editors Alice Weaver and Elizabeth Lesnick for polishing the draft and producing this volume.

Linda Darling-Hammond

As our nation continues the most sustained efforts at school reform it has engaged in for more than a century, it is useful to have the opportunity to rethink the future shape of one of our most important pieces of federal education policy, soon to be reauthorized. Title I of the Elementary and Secondary Education Act of 1965 (now Chapter 1 of the Education Consolidation and Improvement Act of 1981) was enacted to close the gaps between the educational haves and have-nots -- to pursue equity in education by buttressing basic skills instruction for low-income and low-achieving students. Current efforts to restructure schools incorporate these aims but extend them much further, seeking to ensure that all children learn more deeply about more complex matters in more effective ways -- ways that will enable them to meet the demands of a twenty-first century society and workplace in which they will need to be able to analyze, adapt, invent, cooperate, and problem-solve as well as decode and compute.

These demands require major changes in how teaching and learning are managed and assessed, how teachers are prepared and supported, and how schools are organized and run. In brief, they require classrooms that support active learning about complex matters at high levels of performance, and they require schools that support personalized structures for student-adult engagement that can ensure attention to students' individual needs (Lieberman, Darling-Hammond, and Zuckerman, 1991; Elmore, 1990; Sizer, 1984). These demands -- and the realities of life for low-income children at this moment in history -- also require redesign of Chapter 1 in at least five areas:

- Possibilities and incentives for addressing the "whole school" and the "whole child";
- Strategies for achieving greater equity in the allocation of resources to schools serving poor children;
- Approaches to preparing and supporting teachers and other school professionals;
- New models for evaluating programs in line with current understandings of teaching, learning, and assessment;
- Careful assessment of proposals to motivate school improvement.

This paper suggests how changes in Chapter 1 legislation might be pursued in each of these areas and why they are necessary.

Overview of Chapter 1

Over the past quarter century, Chapter 1 has, in many respects, served the nation well -- as it has many of the children who are its direct beneficiaries. In addition to its compensatory education goals, Chapter 1 has influenced many other aspects of equity in schooling. Through its comparability requirements, Chapter 1 helped to equalize intradistrict spending patterns. Its parent advisory councils legitimized the notion of parent involvement in school decision making, and in some places sowed the seeds for new, more participatory governance structures. Its focus on the needs of low-income and low-achieving children created a mandate for change in the way such children are educated, drawing attention and resources to their needs, and probably contributing to a narrowing of the gap in basic-skills achievement between majority and minority students.

Yet some regulations and auditing practices associated with Chapter 1 and other federal categorical programs have encouraged the bureaucratization and compartmentalization of schools. Within Chapter 1, audit requirements to show that funds "supplementing" without "supplanting" normal school expenditures on individual children have led to widespread use of pull-out programs. Across federal categorical programs, the need to categorize children according to various eligibility criteria has often led to segregating or stigmatizing approaches to the delivery of services, as well as to fragmentation of services, which are funded by separate programs through separate offices using separate funding channels. As an unintended consequence, the bureaucratic program structures created by auditing, monitoring, and reporting requirements have fragmented the time and efforts of schoolpeople and children and have deflected resources, or prevented the creative use and merger of resources, for more effective instruction (Kimbrough and Hill, 1981; Soo Hoo, 1990; Klugman, Carter, and Israel, 1979).

In addition, evaluation requirements tied to norm-referenced standardized testing for Chapter 1 have created pressure to "teach to tests," which exerts low-level cognitive demands and often undermines more effective teaching techniques (Darling-Hammond, 1991). Inequalities in access to curriculum and teaching resources, within as well as across schools, further limit the academic opportunities of Chapter 1's intended beneficiaries (Oakes, 1985; Darling-Hammond, 1990a). For these reasons, a rethinking of the forms of federal funding and the nature of regulation for these categorical programs is in order.

Serving the Whole Child in an Improved School Setting

First, since the intended recipients of major federal program services increasingly overlap in population and are concentrated in their geographic locations, "whole-child" and "whole-school" approaches to the funding of compensatory education, along with bilingual education, special education, health, and other auxiliary services, are needed to encourage

more child-centered approaches to schooling and teaching that ensure sensibly integrated help for students. Chapter 1's schoolwide project initiative, launched with the 1988 Hawkins-Stafford Amendments, fits with the current recognition of the school itself as the key unit of change and improvement.

Schoolwide Chapter 1 projects have been noticeably successful in improving teaching, learning, and child services for all students in a number of schools in Philadelphia, New York, and other cities. Among the benefits named by principals of schools that have launched schoolwide projects are the opportunity to serve more students more effectively, to reduce class size and use staff more flexibly, to improve scheduling and coordination of services, to make more resources available to students, and to support professional development and different teaching strategies (Schenck, 1992; U.S. Department of Education, 1992). These projects have enabled students to be supported in more personalized and thoughtful, and less segregated or fragmented, programs while also stimulating school improvement on a broader scale.

This aspect of Chapter 1 should be strengthened and expanded. Eligibility rules should allow more schools access to whole-school project options with necessary safeguards and supports for careful planning and appropriate services. In addition, more "high-impact" schools -- those with many children eligible for various federally funded services, such as bilingual education and special education, as well as Chapter 1 -- should be enabled and encouraged to combine funding from several federal programs. With appropriate safeguards and planning, this would allow students to be better taught in more personalized fashion throughout the day, rather than shuffled from one fragmented pull-out class to another. Finally initiatives to enhance the capacity of schools to combine funding for education, health, and social welfare services for children so as to provide an integrated, workable safety net for children should be aggressively pursued in Chapter 1 and other federal categorical programs.

Among the strategies that Chapter 1 should encourage are intensive staff development programs linked to school improvement initiatives. Also important is the reallocation of professional personnel so that they may work intensively with students for longer periods of time and over several years rather than in compartmentalized programs that split students' time among many adults, none of whom can come to know the child well. In U.S. schools, only 50 percent of professional staff are classroom teachers; the remainder work in specialized offices and roles that serve children at the periphery of the classroom. As a consequence, while staff/pupil ratios in most schools average around 1 to 12, class sizes remain more than twice as high. Students do not receive personalized attention because the organization of the school and the administration of programs focus on specialized divisions of labor rather than on integrated attention to the whole child. Regulatory strategies should be evaluated for the extent to which they enable schools to focus resources on students' needs within their core classrooms, rather than diluting resources into an array of different offices and bureaucracies that fragment students' and teachers' time and that can rarely address the root causes of classroom failure.

Pursuing Greater Equity in Access to Education

Despite some equalization spurred by Chapter 1's comparability requirements, there are still dramatic inequalities in education funding across states, among districts, and among schools within districts (Educational Testing Service, 1991; Kozol, 1991). These inequalities routinely disadvantage students in low-income schools, especially those in central cities and poor rural areas. Inequality is often most obvious in the contrasts between overcrowded, dilapidated schools with large classes lacking equipment and materials in poor neighborhoods and bright, airy facilities where students study in well-equipped small classes in affluent neighborhoods. Less visible but even more pronounced are inequalities in students' access to highly qualified teachers and high-quality curricula. These inequalities start with unequal allocations of funds across and within states; they are compounded by unequal distributions of qualified teachers within districts and further exacerbated by overt and tacit tracking within schools, which rations expert teachers and challenging curricula to the more privileged students (Oakes, 1985, 1986; Darling-Hammond, 1990a; Darling-Hammond and Gendler, in press).

As a consequence of these inequalities, students are placed "at risk" not only from poverty, but also by the schools they attend. Chapter 1 will never be effective at remedying underachievement as long as its services are layered on a system that so poorly educates low-income children to begin with. Current efforts operate from the flawed presumption that "the schools are fine, it's the children who need help." The schools serving large concentrations of Chapter 1 students are frequently not fine, and many of their problems originate with district and state policies and practices that place the schools further at risk.

The recently released interim report of the independent Commission on Chapter 1 observes that "given the inequitable distribution of state and local resources, the current notion that Chapter 1 provides supplemental aid to disadvantaged children added to a level playing field is a fiction" (Commission on Chapter 1, 1992, p. 4). The commission proposes that each state be held accountable for assuring comparability in "vital services" among all its districts as well as in all schools within each district. This proposal should be seriously pursued in the reauthorization of Chapter 1. Among these vital services, perhaps the most important is highly qualified teachers, not just for specific Chapter 1 services but for all classrooms.

Low-income students in many central-city and poor rural schools are routinely taught by a disproportionate number of inexperienced and underprepared teachers, teachers teaching out of their fields of preparation, and a parade of short- and long-term substitutes hired when vacancies cannot be filled. Dramatic differences in salaries and working conditions across school districts explain much of the disparity in teacher supply and qualifications between cities and their generally wealthier suburbs. In recent years, shortages of qualified teachers in subject areas like early childhood education, bilingual education, special education, mathematics, science, and foreign languages have forced cities like New York, Los Angeles,

and Houston to hire thousands of teachers who are not fully prepared. The vast majority of these are assigned to the most disadvantaged central-city schools where working conditions are least attractive and turnover rates are highest (Darling-Hammond, 1990a, 1992).

Furthermore, within schools, inexperienced and underprepared teachers are disproportionately assigned to teach the lowest achieving students. Tracking persists in the face of growing evidence that it does not particularly benefit high achievers and tends to put low achievers at a serious disadvantage (Oakes, 1985, 1986), in part, because teachers are a scarce resource, and thus must be allocated. Evidence suggests that teachers themselves are tracked, with those judged to be the most competent and experienced or with the highest status assigned to the top tracks (Oakes, 1986; Davis, 1986; Findley, 1984; Rosenbaum, 1976). Expert, experienced teachers who are in great demand are rewarded with opportunities to teach the students who already know a lot. New teachers, unprepared teachers, and those teaching outside their field of preparation are often assigned to the students and the classes that nobody else wants to teach, which leaves them practicing on the students who would benefit most from the skills of knowledgeable, experienced teachers.

These circumstances have many unfortunate consequences for the children who are the intended beneficiaries of Chapter 1. Partly as a function of the limited skills and expectations of their teachers (and partly due to the pressures exerted by standardized tests, which are discussed later), students placed in the lowest tracks or in remedial programs often work at a low cognitive level on rote tasks and worksheets that are profoundly disconnected from the skills they need to learn. Rarely are they given the opportunity to talk about what they know, to read real books, or to construct and solve problems in mathematics or science (Oakes, 1985; Davis, 1986; Metz, 1978; Trimble and Sinclair, 1986; Cooper and Sherk, 1989). What these children learn is quite different from what students learn in upper tracks or in schools where good teaching is widespread.

In "Closing the Divide," Robert Dreeben (1987) describes the results of his study of reading instruction and outcomes for 300 black and white first graders across seven schools in the Chicago area. What he found was that differences in reading outcomes among students were almost entirely explained not by socioeconomic status or race but by the quality of instruction the students received:

> Our evidence shows that the level of learning responds strongly to the quality of instruction: having and using enough time, covering a substantial amount of rich curricular material, and matching instruction appropriately to the ability levels of groups. . . . When black and white children of comparable ability experience the same instruction, they do about equally well, and this is true when the instruction is excellent in quality and when it is inadequate. (p. 34)

The study also found, however, that the quality of instruction received by black students was, on average, much lower than that received by white students, thus creating a racial gap in aggregate achievement at the end of first grade. In fact, the highest ability

group in Dreeben's entire sample was in a school in a low-income black neighborhood. These students, though, learned less during first grade than did their lower-aptitude white counterparts, because their teacher was unable to provide the kind of appropriate and challenging instruction this talented group deserved.

Many other studies have also found that differences in school achievement are largely due to the effects of substantially different school opportunities between more and less advantaged students (Barr and Dreeben, 1983; Dreeben and Gamoran, 1986; Dreeben and Barr, 1987; Oakes, 1990; Darling-Hammond and Snyder, 1992). As Oakes (1990) found in looking at the distribution of mathematics and science opportunities across the nation:

> Our evidence lends considerable support to the argument that low-income, minority, and inner-city students have fewer opportunities . . . . They have considerably less access to science and mathematics knowledge at school, fewer material resources, less-engaging learning activities in their classrooms, and less-qualified teachers. . . . Moreover, our findings are likely to be equally relevant for subject areas other than mathematics and science. The differences we have observed are likely to reflect more general patterns of educational inequality. (pp. x-xi)

The differences that teacher preparation makes for student learning are often apparent at the individual level. When school staffing patterns create substantial imbalances in teacher expertise across schools, the effects are startling. When Armour-Thomas and colleagues (1989) compared a group of exceptionally effective elementary schools with a group of low-achieving schools with similar demographic characteristics in New York City, they found that differences in teacher qualifications and experience accounted for roughly 90 percent of the variance in student reading and mathematics scores at grades 3, 6, and 8. Far more than any other factor, teacher expertise made the difference in what children learned.

If, in fact, the interaction between teachers and students is the most important aspect of effective teaching, then reducing inequality in learning has to rely on policies that provide equal access to competent, well-supported teachers. The comparability requirements for Chapter 1 should take into account the extent to which states and school districts ensure that "vital services" -- including experienced, qualified teachers, along with challenging curriculum opportunities and up-to-date curriculum materials and equipment -- are equally distributed among schools. In addition, improving instruction for low-income students will require improving the capacity of all teachers to understand how children learn and develop, how a variety of curricular and instructional strategies can address their needs, and how changes in school and classroom organization can support children's growth and achievement.

Teacher Preparation and Staff Development

As noted above, policy makers have nearly always answered the problem of teacher shortages in cities by lowering standards so that people with little or no preparation for teaching can be hired. Although this practice is often excused by the presumption that virtually anyone can figure out how to teach, a number of reviews of research summarizing the results of more than 100 studies have concluded that fully prepared and certified teachers are more successful with students than are teachers without full preparation (Darling-Hammond, 1992; Evertson, Hawley, and Zlotnik, 1985; Ashton and Crocker, 1986; Ashton and Crocker, 1987; Greenberg, 1983; Druva and Anderson, 1983).

The extent and kind of teacher preparation are especially important in determining the effectiveness of teachers in "school-based" subjects (those subjects students tend to learn primarily in school rather than through informal learning outside of school), such as mathematics, science, and early reading. Teacher training is also a critical determinant of the use of teaching strategies that encourage higher-order learning and the use of strategies responsive to students' needs and learning styles.

Strickland (1985) stresses that, for early literacy development, elementary school teacher preparation programs must help teachers develop programs able to accommodate a variety of cognitive styles and learning rates, with activities that broaden rather than reduce the range of possibilities for learning. They should prepare teachers to understand how children acquire and develop language skills as well as the nature of children's growth and development. These understandings should undergird a knowledge of appropriate methods for fostering language growth at various stages of development. Comer (1988) also emphasizes the importance of preparing teachers with a strong background in child development as a key to the kind of teaching that has been so successful in his programs. The evidence clearly indicates that such preparation makes a difference in what children learn. Yet very few teachers have received this kind of preparation, and even fewer of them are teaching in schools in high-poverty areas. Significant funding should be provided through Chapter 1, perhaps as a proportion of state and local education agency grants, for staff development for teachers and administrators. Particularly in schools with high concentrations of Chapter 1 students (and students in other federal programs), funding should be used for schoolwide professional development programs that are aimed at improving the general caliber of teaching as a *preventive* to remediation.

In addition, funding strategies for targeted teacher preparation should be developed to ensure an adequate supply of highly qualified teachers to high-impact Chapter 1 schools. These strategies can borrow from the teacher preparation initiatives that accompanied the Education for All Handicapped Children Act (PL 94-142), which helped to ensure that high-quality teacher education programs and qualified teachers would become available, and those used by the Urban Teachers Corps during the 1970s to encourage recruits to teach in urban schools. Such strategies would include grants to teacher education programs, especially in

urban areas, to develop strong programs to prepare teachers in up-to-date methods for the new teaching, learning, and assessment demands posed by recent reforms, along with heightened understanding of the needs of "at-risk" learners. These should be accompanied by forgivable loan programs to encourage recruits to engage in rigorous preparation and then commit themselves to teaching in high poverty urban and rural schools.

New Evaluation Strategies and Standards

In recent years, there has been a growing consensus among researchers, practitioners, and policy makers that the nature of assessment in American schools must change in fundamental ways. Current testing practices are poor measures of what students know and can do, and they undermine instruction focused on the acquisition of higher-order thinking skills and complex performance abilities, contributing to an underachieving curriculum in American schools (Boyer, 1983; Goodlad, 1984; McKnight *et al.*, 1987; Madaus, West, Harmon, Lomax, and Viator, 1992). The requirement that Chapter 1 programs be evaluated using nationally normed standardized tests contributes to the overuse of such tests, especially with very young students, for whom they are least reliable and valid. Use of mass-administered, multiple-choice, standardized tests as the sole measure for identifying young students for Chapter 1 services is now a widespread practice, though it is one not endorsed by professional psychologists, testing experts, researchers, or early childhood educators (National Association for the Education of Young Children [NAEYC], 1988, 1991). This Chapter 1 requirement has impeded state and local efforts to reform their assessment practices and is widely viewed as in need of revision.

The Problem with Current Tests

In contrast to testing in most other countries, American testing is dominated by norm-referenced multiple-choice instruments designed to rank students cheaply and efficiently. Initially created to make tracking and sorting of students more efficient, these instruments were not intended to support or enhance instruction. Because of the way the tests are constructed, they exclude a great many kinds of knowledge and types of performance we expect from students, placing test-takers in a passive, reactive role, rather than one that engages their capacities to structure tasks, produce ideas, and solve problems (National Research Council [NRC], 1982). The tests are poor measures of students' actual performance abilities and inappropriate tools for many of the purposes they are used to serve (NRC, 1982; Darling-Hammond, 1991).

Most traditional standardized tests do not reflect current understandings of how students learn. They are based on an outmoded theory of learning as the accumulation and recall of isolated facts and skills. They do not reflect present knowledge that people learn in meaningful and purposeful contexts by connecting what they already know with what they

are trying to learn (Gardner, 1983; Kantrowitz and Wingert, 1989; Resnick, 1987). "Thinking skills" are the foundation for building "basic skills," not the other way around, as many testing programs assume. Furthermore, real skills must be demonstrated in complex performance contexts, not on tasks demanding only recognition of discrete facts. By focusing on rote skills rather than conceptual learning, the tests often miss the forest for the trees.

Because most traditional standardized tests provide only a limited measure of a narrow aspect of learning or development, they are poor predictors of how students will perform in other settings, and they are unable to provide information about why students score as they do. Because they record only final answers and report only numerical scores, mass-administered standardized tests do not provide information about how children tackle different tasks or what abilities they rely on in their problem solving. This promotes a view of children as having deficits that need to be remediated rather than as having individual approaches to learning and unique strengths that can be supported and developed. It also fails to provide enough information about areas of difficulty to inform instructional strategies for addressing them (NAEYC, 1988; Bredekamp and Shepard, 1989).

These shortcomings of American tests have become more problematic as test scores have been used to make important educational decisions. Schools' efforts to teach to the tests have reduced students' opportunities for higher-order learning, because classwork oriented toward recognizing the answers to multiple-choice questions does not heighten students' proficiency in aspects of subjects not tested, such as analysis, complex problem solving, and written and oral expression. Many studies have found that because of test-oriented teaching, American students listen, read textbook sections, respond briefly to questions, and take short-answer and multiple-choice quizzes. They rarely plan or initiate anything, create their own products, read or write anything substantial, engage in analytic discussions or in projects requiring research, invention, or problem solving (National Assessment of Educational Progress, 1981; Boyer, 1983; Goodlad, 1984; Darling-Hammond and Wise, 1985; McKnight *et al.*, 1987).

The results can be seen in U.S. achievement trends. Since about 1970, basic skills test scores have increased slightly while scores on assessments of higher-order thinking have been steadily declining in virtually all subject areas. Officials of the National Assessment of Educational Progress (NAEP), the National Research Council (NRC) and the National Councils of Teachers of English and Mathematics, among others, have all attributed this decline to schools' emphasis on tests of basic skills (NAEP, 1979; NRC, 1979; Office of Productivity, Technology, and Innovation, 1980). In addition, the emphasis on rote learning contributes to American students' consistently dismal rankings on international achievement tests (McKnight *et al.*, 1987).

As one of the NAEP reading assessments found: "Only 5 to 10 percent of students can move beyond initial readings of a test; most seem genuinely puzzled at requests to explain or defend their points of view." This report explained that current methods of testing

reading require short responses and lower-level cognitive thinking, resulting in "an emphasis on shallow and superficial opinions at the expense of reasoned and disciplined thought . . . , [thus] it is not surprising that students fail to develop more comprehensive thinking and analytic skills" (NAEP, 1981).

Unfortunately, the effects of test misuse have often been most damaging for the students these tests were intended to help. Many studies have found that students placed in the lowest tracks or in compensatory programs are most apt to experience instruction geared only to multiple-choice tests, working at a low cognitive level on test-oriented tasks that do not resemble or support the development of real performance skills. Because the content and format of the tests are so far removed from the ways actual performance is exhibited, teachers must drill students on test skills, such as recognizing and marking multiple-choice answers, rather than real skills, such as reading books and talking or writing about what they mean. In short, students are denied the opportunity to develop the capacities they will need for the future, in large part because our tests are so firmly pointed at educational goals of the past.

Improving Assessment

Recognizing these problems, many schools, districts, and states have recently begun to develop different forms of assessment for students. States like Vermont, California, Connecticut, Maryland, and New York, along with many districts across the country, are developing approaches much like the assessment systems that prevail in other countries around the world. These include essay examinations, scientific experiments, and exhibitions in areas like debating and the arts. They also include portfolios of students' work and projects that require analysis, investigation, experimentation, cooperation, and written, oral, or graphic presentation of findings. These assessments require students to think analytically and demonstrate their proficiency as they would in real-life performance situations.

These initiatives will falter or flourish depending on the directions taken by federal policy makers. It is critically important that Chapter 1 evaluation standards not require the continued use of testing instruments that exhibit limited validity and that, by their influence on instruction, severely limit the kinds of learning experiences to which students are exposed. School districts and states should be encouraged to propose alternative forms of assessment that reliably assess students' performance capabilities, including performance tasks, evaluations of student work, and systematic teacher observations of students.[1]

There is a wide range of assessment instruments available and in development that can be used to diagnose student needs and to evaluate student progress. These include

1. Some existing systems, such as the Primary Language Record (PLR) developed in England and used in an increasing number of U.S. schools, provide a framework for combining all of these data sources in the evaluation of early literacy acquisition for children. The PLR also provides data in the form of scale scores that can be maintained longitudinally and reported in aggregated form.

individually administered written, oral, and performance assessments, some of which are normed and others of which are criterion-referenced, along with observation systems and methods for evaluating student work, behavior, and demonstrated abilities over time.

Schools and districts should be required to show that they have an assessment *system* for evaluating student progress and program accomplishments that

- Measures progress toward the full range of program goals for students;
- Is as authentic as possible (i.e., represents real performance tasks and situations);
- Measures higher-order skills and performance abilities;
- Provides multiple methods and occasions for students to demonstrate their skills and knowledge, including many different kinds of performance and multiple assessment strategies and measures;
- Allows teachers to evaluate student growth in a longitudinal, cumulative fashion using several kinds of evidence (e.g., samples of work, observations, performance on tasks);
- Allows assessment of student progress using longitudinal gains on criterion-referenced measures for the individual students in the program.

The last of these criteria suggests an essential change in how programs are evaluated. As discussed below, current evaluations that look at program "effectiveness" using cross-sectional measures of norm-referenced achievement produce spurious results because their methods obscure both who and what are being measured. Program populations change from the time one set of average test scores is produced until the next; thus, a stable population is not being assessed, and "gains" or "losses" may be a function of changes in who is taking the test, rather than in how individual students score. In high-mobility urban schools, these changes in school or program population are frequently substantial, so that few of the original students remain when subsequent measures are taken and those who are tested were not there at the previous testing time.

At the same time, norm-referenced tests, because of the way they are constructed, cannot allow for real evaluations of progress for individual students. Test items are selected to ensure that items have "discriminatory power" so that students may be dispersed along a normal curve. Items to which too many, too few, or the wrong subset of students know the answer are eliminated, regardless of their educational significance. This skews test content in a variety of ways, making it difficult to evaluate what students have actually learned. In addition, 50 percent of students must always score below the mean. Thus, a clear measure of how much and what kind of progress students have actually achieved, either individually

or collectively, is never possible. The standard is how students rank against each other, not how any of them measure up against a stable standard of performance.

These shortcomings have contributed to a widespread perception that local Chapter 1 evaluations are essentially meaningless. Because of the instability of the normal curve equivalent (NCE) scores currently used to identify Chapter 1 schools for program improvement interventions, about half of the schools so identified have "tested out" of the improvement program before ever implementing a plan. Those that did implement plans focused more on raising average scores than on diagnosing program flaws or changing practices (U.S. Department of Education, 1992).

One rationale for requiring evidence based on nationally normed tests is to gain comparable nationwide data for evaluating the entire Chapter 1 program. This purpose turns out to be poorly served by the data districts collect on norm-referenced pre- and post-tests (U.S. Department of Education, 1992). For national program evaluation purposes, data on individual students are not necessary, and current practices are actually counterproductive. When Chapter 1 requires evaluation, much more carefully designed sample studies will always be needed to provide the controls required for drawing reasonable inferences about program effects. Data on how Chapter 1 schools are performing can also be attained through matrix sampling techniques that do not require testing every student.

These evaluations do not need to depend on norm-referenced examinations: Criterion-referenced assessments can produce comparable data by their incorporation into appropriately fashioned research.[2] Meanwhile, local assessment strategies should be developed to better serve local needs for informing instruction and for providing useful information to parents. Individual student assessments should not be used for the purposes of program evaluation. This use limits the selection of assessment tools to the least instructionally useful and distorts their results in order to try to serve purposes they can never, in any event, serve well.

These evaluation considerations are also linked to the intended and unintended incentives that influence who is to be served by the programs and how the programs serve

---

2. One nationally administered criterion-referenced assessment is the National Assessment of Educational Progress (NAEP). Though it is conducted on a cross-sectional sampling basis and thus is not appropriate for charting the progress of individual students who have been in Chapter 1 or other programs, NAEP can, nonetheless, give information on how students in Chapter 1 schools are performing in states across the country. National longitudinal studies, such as the National Educational Longitudinal Study (NELS), can follow the progress of samples of students over many years.

In addition, nearly all states have developed criterion-referenced assessments of student performance that can provide information that is at least as useful as that provided by commercially developed norm-referenced tests. In many cases, these state assessments are becoming increasingly sophisticated and performance-oriented, capable of evaluating a broader range of skills and abilities than other standardized tests do. Carefully designed studies can incorporate such information for samples of Chapter 1 and similar non-Chapter 1 students tracked over a period of time to assess progress in learning.

them. These incentives, and recent proposals for altering them, are discussed in the next section.

Incentives for School Improvement

There are at least three aspects of Chapter 1 operations that exert intended or unintended incentives for how schools serve students: (1) how funds are allocated to schools and students; (2) how programs are evaluated; and (3) what efforts are made to encourage program improvement.

Chapter 1 has been criticized for providing counterproductive incentives for serving students well. In its interim report, the ***National Assessment*** of Chapter 1 quotes mathematics teacher Jaime Escalante as saying:

> When I went to one administrator to ask for the textbooks I felt I needed to raise students' scores to even the minimum standard of competency in the routine State and District mandatory testing, I was told: "No, Jaime, those books do not qualify for Title I. They are clearly not remedial." This same individual also told me, in complete seriousness: "Jaime, if you get better test scores for these kids then we're not going to qualify for Title I money." (U.S. Department of Education, 1992, p. 69)

The disincentive Escalante refers to is the potential loss of funds for schools whose students improve their performance. This disincentive is caused by the requirement that districts allocate their Chapter 1 monies according to both poverty and educational need. "Educational need" is generally measured by the proportion of students scoring below some cutoff point on a standardized test. Obviously, when districts allocate funds based on the number of low-scoring students in each school, schools lose money as their students score higher. The incentive is not to be too effective at improving student achievement. The most straightforward approach to remedying this disincentive, and arguably the one most faithful to the original intent of Title I, is to allocate funds to schools strictly on the basis of the concentration of children in poverty, with even more emphasis than currently exists on targeting funds to schools with high concentrations of poverty.

This would ensure that schools serving low-income children can build a resource-rich environment for students who have fewer material resources at home. If these schools then produce highly successful students with these greater resources, the fundamental goals of Chapter 1 will have been served, and the funding that supports that success should be continued, rather than reduced.

This solution is much more reasonable than that suggested by other proposals that are

currently being put forth to create a reward-and-sanction system of "incentives" for Chapter 1. For example, the Commission on Chapter 1 proposes to create "an outcome-based system of accountability [that] will also form the basis for new plans of enforcement that utilize both incentives and sanctions" (1992, p. 4). Using new performance standards and assessments (based on new "curricula frameworks" not yet defined or developed), this approach to accountability would reward high-scoring schools with greater resources, and would penalize low-scoring schools with sanctions that are as yet unspecified.

Such a system is certain to create other counterproductive incentives of its own, based on what has happened in schools and districts that have followed similar courses of action. When program "outcomes" are the focus of reward-and-sanction systems, incentives are created for manipulating the population of students in the school or program and taking the tests. A purely outcomes-based accountability system confuses the quality of education offered by schools or programs with the changing needs of the students they enroll. Such a system can ultimately restrict opportunities by discouraging schools from serving the most educationally needy students.

Research has already found that efforts to use "outcome" measures like average test scores to reward and punish schools have resulted in a number of practices that harm students and fail to improve programs: for example, labeling large numbers of low-scoring students for special education placements so that their scores will not "count" in school reports, retaining students in grade so that their relative standing will look better on "grade-equivalent" scores, excluding low-scoring students from admissions to schools or programs, and encouraging such students to transfer to other schools or drop out (Chancellor's Commission on Minimum Standards, 1988; Darling-Hammond, 1991; Koretz, 1988; Shepard and Smith, 1986; Smith and Shepard, 1988; Smith *et al.*, 1986).

Equally important, these policies further exacerbate existing incentives for talented staff to opt for school placements where students are easy to teach, and school stability is high. Capable staff cannot be expected to risk losing rewards or incurring sanctions by volunteering to teach where many students have special needs and performance standards will be more difficult to attain. Such an approach further compromises the educational chances of disadvantaged students, who are already served by a disproportionate share of those teachers who are inexperienced, unprepared, and underqualified (Oakes, 1990; Darling-Hammond, 1990a, 1992).

It is critical to understand that penalties administered to schools ultimately penalize the students who attend those schools. Whether the penalty is some form of probation or "deregistration" (that is, removal of state or professional approval to operate) or sanctions for staff individually or collectively, this approach to school improvement will exacerbate a downward spiral for many of the most troubled schools by making it even more difficult for them to recruit qualified staff. Promises of bonuses for "successful" schools, coupled with threats of penalties if scores decrease, will also reduce the likelihood that high-performing schools of choice will admit the most educationally needy Chapter 1 students. Unlike

private-sector businesses that produce inanimate products, schools cannot be treated as independent of the students they serve. Schools are collections of students as well as collections of programs, and they both influence and are influenced by the populations of students they serve. If short-sighted incentive schemes make it risky to serve the most educationally needy students in Chapter 1 programs, these students will be pushed out or kept out of programs from which they could benefit.

The end result of this approach to accountability is more statistical game playing and greater harm to educationally disadvantaged students, but not higher levels of learning for them or for the student population as a whole. It is virtually certain that a rewards-and-sanctions approach to "accountability" for Chapter 1 would result in many low-scoring students being inappropriately shifted to special education programs, being retained in grade so that their scores look better (a practice that has been shown to undermine both long-range achievement and graduation rates),[3] and being excluded from schools of choice, which promise to be an increasing number of public schools. These results are predictable whether the measures are traditional standardized tests or new assessment measures, and whether the statistic used is the percentage of students attaining a particular score or the presumed "gain" in performance for the aggregated scores of program students at two points in time.

Accountability that supports responsible practice can occur only when a useful set of processes exists for guiding practice and for using information to improve practice (Darling-Hammond and Ascher, 1991). Thus, assessments of student performance provide information for an accountability system; they are not the system itself. Outcome standards alone cannot guarantee accountable schools. In fact, if these are improperly designed or used, they can actually undermine accountability.

There are two basic problems with the approaches to outcomes-based accountability that have already been tried in some states and school districts. First, the measurement of "outcomes" or "gains" has been seriously flawed. Second, the basic theory of the incentives needed for school improvement is inadequate.

As noted earlier, evaluations that look at program "effectiveness" using aggregated cross-sectional achievement measures produce spurious results unless the same students are in the program at the two points in time when measures are taken. This is almost never the case. The problem is particularly pronounced in the many high-poverty Chapter 1 schools with high transiency rates, some of which have pupil turnover rates exceeding 100 percent

3. When students who have been retained in grade are compared with students of equal achievement levels who have been promoted, the students who were retained are consistently behind on both achievement and social-emotional measures (Shephard and Smith 1986; Holmes and Matthews 1984; Rose, Medway, Cantrell, and Marus, 1983). Furthermore, research suggests that being retained in grade increases the likelihood of dropping out by 40 to 50 percent. A second retention increases the risk by 90 percent (Mann 1987; see also Carnegie Council on Adolescent Development 1989; Wehlage, Rutter, Smith, Lesko, and Fernandez, 1990; Massachusetts Advocacy Center, 1988).

annually.[4] Furthermore, given the wide array of differences among students and programs, it is not sensible to create a single expectation for the magnitude of "progress" each program should achieve. While some schools serve students only marginally in need of extra help, others serve students who are homeless and frequently absent, students who are recent immigrants with no English and little education prior to their arrival, or students who are children of drug-abusing parents, sometimes badly injured by prenatal drug use and in need of intensive assistance to make any progress at all.

To make any inference about program contributions to student learning, individual data for only those students who have been in the program over the time period assessed should be used. Furthermore, these data should be longitudinal, criterion-referenced, and multifaceted, taking into account the special needs and characteristics of pupils served, along with several different kinds of performance measures. Student data should be coupled with information about the nature and appropriateness of program practices. Judgments of program quality must be made on the basis of careful, sophisticated assessments of individual student progress along with professionally sound assessments of program services and practices.

Even with more thoughtful and careful assessments of program quality, such as those described above, the question of how to stimulate improvements in low-quality programs deserves more serious attention than it has thus far received. The notion that rewards and sanctions to schools will result in improvements is based on two assumptions: (1) that the basic problem in poor programs is lack of staff motivation, which can be cured by a system of rewards and penalties; and (2) that the sources of any problems are created at the school level.

Both of these assumptions are highly questionable. In general, failure to implement educational change is a result of lack of knowledge about how and what to change, along with lack of capacity to implement changes (Fullan, 1992; Darling-Hammond, 1990b). In addition, many problems of low capacity and expertise within schools are created by educational agencies and policies outside the school and can be solved only by engaging these external agents (Oakes, 1990; Darling-Hammond, 1990b, 1992). Finally, as already discussed, such reward-and-sanction schemes create unintended incentives to avoid serving the most educationally needy students.

A useful approach for Chapter 1 accountability would maintain the intentions of the Hawkins-Stafford Amendments to stimulate program improvement with thoughtful diagnosis, staff development, and schoolwide change initiatives, but would rely on more careful methods of identifying schools for program improvement, using measures of both practices and student learning (as discussed above). Program-improvement funds should be used to

4. This occurs when there are more student entrances and departures during a single school year -- individual students leaving and returning more than once or many students enrolling for only a short time before leaving -- than the total number of students who are counted as the base student population.

enable outside teams of highly expert reviewers to work with school staff in evaluating their practices and developing a plan. These funds would also support the schoolwide staff development needed to implement new, more effective practices. Diagnosis should include assessment of those district or state-imposed factors that may be undermining school performance, including the availability of resources and qualified personnel. Remedying such factors should be part of the district and state responsibility in overseeing and supporting the plan's implementation.

Chapter 1 can be a critical linchpin in ensuring that low-income children benefit from the efforts of the nation's school reform movement. This role will be strengthened as the program

1. Continues and expands its emphasis on whole school/whole child approaches to education and related services;

2. Creates more leverage for equalizing access to educational resources;

3. Extends its capacity for supporting teacher and school learning and development;

4. Supports more sophisticated approaches to program evaluation and student assessment;

5. Develops strategies for school funding and school improvement that enhance school capacity rather than creating disincentives for serving educationally needy students well.

References

Armour-Thomas, E., Clay, C., Domanico, R., Bruno, K., and Allen, B. (1989). ***An Outlier Study of Elementary and Middle Schools in New York City: Final Report.*** New York: New York City Board of Education.

Ashton, P., and Crocker, L. (1986). "Does Teacher Certification Make a Difference?" ***Florida Journal of Teacher Education, 3***: 73-83.

Ashton, P., and Crocker, L. (1987). "Systematic Study of Planned Variations: The Essential Focus of Teacher Education Reform." ***Journal of Teacher Education, 38***(3): 2-8.

Barr, R., and Dreeben, R. (1983). ***How Schools Work.*** Chicago: University of Chicago Press.

Boyer, E. (1983). ***High School.*** New York: Harper and Row.

Bredekamp, S., and Shepard, L. (1989). "How Best to Protect Children from Inappropriate School Expectations, Practices and Policies." ***Young Children, 44***(3): 14-24.

Carnegie Council on Adolescent Development (1989). ***Turning Points: Preparing Youth for the 21st Century.*** New York: Carnegie Corporation of New York.

Chancellor's Commission on Minimum Standards (1988). ***Contextual Factors Influencing the Setting and Implementation of Minimum Standards.*** Report to the Chancellor from the New York City Chancellor's Commission on Minimum Standards.

Comer, J. P. (1988). "Educating Poor Minority Children." ***Scientific American, 259***(5): 42-48.

Commission on Chapter 1 (1992). ***Commission on Chapter 1: Interim Report.*** Mimeograph.

Cooper, E., and Sherk, J. (1989). "Addressing Urban School Reform: Issues and Alliances." ***Journal of Negro Education, 58***(3): 315-331.

Darling-Hammond, L. (1990a). "Teacher Quality and Equality." In J. Goodlad and P. Keating (Eds.), ***Access to Knowledge: An Agenda for Our Nation's Schools.*** New York: College Board.

Darling-Hammond, L. (1990b). "Achieving Our Goals: Superficial or Structural Reforms?" ***Phi Delta Kappan, 72***(4): 286-295.

Darling-Hammond, L. (1991). "Implications of Testing Policy for Educational Quality and Equality." ***Phi Delta Kappan, 73***(3): 220-225.

Darling-Hammond, L. (1992). "Teaching and Knowledge: Policy Implications of Alternate Routes to Teacher Certification." ***Peabody Journal of Education***, *67*(3): 123-154.

Darling-Hammond, L., and Ascher, C. (1991). ***Creating Accountability in Big City School Systems***. New York: National Center for Restructuring Education, Schools, and Teaching and ERIC Clearinghouse on Urban Education.

Darling-Hammond, L., and Gendler, T. (in press). "Inequality and Access to Knowledge." In J. Banks (Ed.), ***Handbook of Multicultural Education***. New York: Macmillan.

Darling-Hammond, L., and Snyder, J. (1992). "Traditions of Curriculum Inquiry: The Scientific Tradition." In P. Jackson (Ed.), ***Handbook of Research on Curriculum***. New York: Macmillan.

Darling-Hammond, L., and Wise, A. E. (1985). "Beyond Standardization: State Standards and School Improvement." ***The Elementary School Journal***, *85*(3): 315-336.

Davis, D. G. (1986). ***A Pilot Study to Assess Equity in Selected Curricular Offerings across Three Diverse Schools in a Large Urban School District: A Search for Methodology***. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Druva, C. A., and Anderson, R. D. (1983). "Science Teacher Characteristics by Teacher Behavior and by Student Outcome: A Meta-Analysis of Research." ***Journal of Research in Science Teaching***, *20*(5): 467-479.

Dreeben, R. (1987). "Closing the Divide: What Teachers and Administrators Can Do to Help Black Students Reach Their Reading Potential." ***American Educator***, *11*(4): 28-35.

Dreeben, R., and Gamoran, A. (1986). "Race, Instruction, and Learning." ***American Sociological Review***, *51*(5): 660-669.

Dreeben, R., and Barr, R. (1987). ***Class Composition and the Design of Instruction***. Paper presented at the Annual Meeting of the American Education Research Association, Washington, DC:

Educational Testing Service (1991). ***The State of Inequality***. Princeton: Author.

Elmore, R. F. (Ed.) (1990). ***Restructuring Schools: The Next Generation of Educational Reform***. San Francisco: Jossey-Bass.

Evertson, C., Hawley, W., and, Zlotnick, M. (1985). "Making a Difference in Educational Quality through Teacher Education." ***Journal of Teacher Education***, *36*(3): 2-12.

Finley, M. K. (1984). "Teachers and Tracking in a Comprehensive High School." *Sociology of Education*, *57*(4): 233-243.

Fullan, M. (1992). ***The New Meaning of Educational Change***. New York: Teachers College Press.

Gardner, H. (1983). ***Frames of Mind: The Theory of Multiple Intelligences***. New York: Basic Books.

Goodlad, J. (1984). ***A Place Called School: Prospects for the Future***. New York: McGraw-Hill.

Greenberg, J. D. (1983). "The Case for Teacher Education: Open and Shut." ***Journal of Teacher Education***, *34*(4): 2-5.

Holmes, C. T., and Matthews, K. M. (1984). "The Effects of Nonpromotion on Elementary and Junior High School Pupils: A Meta-Analysis." ***Review of Educational Research***, *54*(2): 225-236.

Kantrowitz, B., and Wingert, P. (1989). "How Kids Learn." ***Newsweek***, April 17, 1989, pp. 50-57.

Kimbrough, J., and Hill, P. T. (1981). ***The Aggregate Effects of Federal Education Programs***. Santa Monica: RAND Corporation.

Klugman, E., Carter, S., and Israel, S. (1979). ***Too Many Faces: A Study of Teacher Fragmentation in the Elementary School***. Boston: Wheelock-Malden Teacher Corps Collaborative.

Koretz, D. (1988). "Arriving in Lake Wobegon: Are Standardized Tests Exaggerating Achievement and Distorting Instruction?" ***American Educator***, 12(2): 8-15, 46-52.

Kozol, J. (1991). ***Savage Inequalities***. New York: Crown Publications.

Lieberman, A., Darling-Hammond, L., and Zuckerman, D. (1991). ***Early Lessons in Restructuring Schools***. New York: National Center for Restructuring Education, Schools, and Teaching.

Madaus, G. F., West, M., Harmon, M., Lomax, R., and Viator, K. (1992). ***The Influence of Testing on Teaching Math and Science in Grades 4-12***. Chestnut Hill, MA: Boston College Center for the Study of Testing, Evaluation, and Educational Policy.

Mann, D. (1987). "Can We Help Dropouts? Thinking about the Undoable." In G. Natriello (Ed.), ***School Dropouts: Patterns and Policies***. New York: Teachers College Press.

Massachusetts Advocacy Center and the Center for Early Adolescence (1988). ***Before It's Too Late: Dropout Prevention in the Middle Grades***. Boston: Author.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, S. O., Travers, K. J., and Cooney, T. J. (1987). ***The Underachieving Curriculum: Assessing U.S. School Mathematics from an International Perspective***. Champaign, IL: Stipes Publishing.

Metz, M. H. (1978). ***Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools***. Berkeley: University of California Press.

National Assessment of Educational Progress (NAEP) (1979). ***Changes in Mathematical Achievement, 1973-78***. Denver: Author.

National Assessment of Educational Progress (NAEP) (1981). ***Reading, Thinking, and Writing: Results from the 1979-80 National Assessment of Reading and Literature***. Denver: Author.

National Association for the Education of Young Children (NAEYC) (1988). "NAEYC Position Statement on Developmentally Appropriate Practice in the Primary Grades, Serving 5 through 8 Year Olds." ***Young Children***, ***43***(2): 64-84.

National Association for the Education of Young Children. (1991). "Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 through 8." ***Young Children***, ***46***(3): 21-38.

National Research Council (NRC) (1979). ***The State of School Science***. Washington, DC: Commission on Human Resources.

National Research Council (NRC) (1982). ***Ability Testing: Uses, Consequences, and Controversies***. A. K. Wigdor, and R. Garner (Eds.). Washington, DC: National Academy Press.

Oakes, J. (1985). ***Keeping Track: How Schools Structure Inequality***. New Haven: Yale University Press.

Oakes, J. (1986). "Tracking in Secondary Schools: A Contextual Perspective." ***Educational Psychologist***, ***22***(2): 129-154.

Oakes, J. (1990). ***Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science***. Santa Monica: RAND Corporation.

Office of Productivity, Technology, and Innovation (1980). ***Learning Environments for Innovation***. Washington, DC: Department of Commerce.

Resnick, L. B. (1987). ***Education and Learning to Think***. Washington, DC: National

Academy Press.

Rose, J. S., Medway, F. J., Cantrell, V. L., and Marus, S. H. (1983). "A Fresh Look at the Retention-Promotion Controversy." ***Journal of School Psychology***, ***21***(3): 201-211.

Rosenbaum, J. E. (1976). ***Making Inequality: The Hidden Curriculum of High School Tracking***. New York: John Wiley & Sons.

Schenck, A. (1992). "Schoolwide Project Survey: School Building Questionnaire." Hampton, NH: RMC Research Corp. (unpublished tabulations).

Shephard, L. A., and Smith, M. L. (1986). "Synthesis of Research on School Readiness and Kindergarten Retention." ***Educational Leadership***, ***44***(3): 78-86.

Sizer, T. (1984). ***Horace's Compromise***. Boston: Houghton Mifflin.

Smith, F. *et al.* (1986). ***High School Admission and the Improvement of Schooling***. New York: New York City Board of Education.

Smith, M. L., and Shepard, L. A. (1988). "Kindergarten Readiness and Retention: A Qualitative Study of Teachers' Beliefs and Practices." ***American Educational Research Journal***, ***25***(3): 307-333.

Soo Hoo, S. (1990). "School Renewal: Taking Responsibility for Providing an Education of Value." In J. Goodlad and P. Keating (Eds.), ***Access to Knowledge: An Agenda for Our Nation's Schools***. New York: College Board.

Strickland, D. (1985). "Early Childhood Development and Reading Instruction." In C. Brooks (Ed.), ***Tapping Potential: English and Language Arts for the Black Learner***. Urbana, IL: National Council of Teachers of English.

Trimble, K., and Sinclair, R. L. (1986). ***Ability Grouping and Differing Conditions for Learning: An Analysis of Content and Instruction in Ability-Grouped Classes.*** Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

U.S. Department of Education (1992). ***National Assessment of the Chapter 1 Program: The Interim Report***. Washington, DC: Author.

Wehlage, G. G., Rutter, R. A., Smith, G. A., Lesko, N., and Fernandez, R. R. (1990). ***Reducing the Risk: Schools as Communities of Support***. New York: The Falmer Press.

This publication was printed on recycled papers.