

Students with Disabilities and Accountability Reform: Findings from the New York Case Study

Katherine M. Nagle
University of Maryland

Margaret J. McLaughlin
University of Maryland

Victor Nolet
Western Washington University

Kimber Malmgren
University of Wisconsin, Madison

Educational Policy Reform Research Institute

The Institute for the Study of Exceptional Children and Youth
University of Maryland
1308 Benjamin Building
College Park, Maryland 20742-1161

March, 2007

Funding for this research work was provided by the U.S. Department of Education, Office of Special Education Programs (Grant #H324P000004). Opinions expressed in this paper are those of the authors, and do not necessarily reflect the views of the U.S. Department of Education or the Office of Special Education Programs.

This document is one of four individual case study reports presenting the qualitative findings from a five-year investigation of the impact of accountability reform on students with disabilities in four states, eight districts, and twenty schools. During the late 1980s and 1990s, many states passed legislation requiring school systems to be more accountable for students' learning. However, the impetus for the most recent changes in state accountability systems originated in state responses to federal concerns about the quality and equity of public education. These concerns can be considered as the next phase of the standards-based reform movement.

The standards-based reform movement, which began during the 1970s, has been integrated into federal education policy, most notably the last two reauthorizations of the Elementary and Secondary Education Act (ESEA), and has become increasingly focused on holding schools accountable for the academic performance of all their students. The drive for greater accountability and educational equity is embedded in the 2001 reauthorization of ESEA, known as the No Child Left Behind Act (NCLB). NCLB aims to increase student achievement, improve schools, provide parents and the community with better information, and close some long-lasting and troubling achievement gaps between disadvantaged students and their peers (Cohen, 2002).

The new requirements for standards and assessments are rigorous and largely build on the existing Title I requirements promulgated under the Improve America's Schools Act. For most states however, *performance-based* accountability requires them to chart a course into new and unfamiliar territory. Particularly controversial is the requirement that the *same* annual academic achievement objectives be determined, met,

and reported for specific subgroups, including most students with disabilities, racial/ethnic minority groups, and students with limited English proficiency (LEP).

This case study report presents qualitative data, collected over a two-year period—2001 and 2002—from two levels of the New York public education systems on the impact of accountability reform on students with disabilities. The following three research questions guided this component of the project:

- 1) How do broad education policies that incorporate high-stakes accountability include consideration of students with disabilities?
- 2) How are students with disabilities affected by education accountability reforms?
- 3) What impact have educational accountability mechanisms had on students with disabilities at the system and individual student level?

The remainder of this report is divided into five sections. The first section contains information regarding the methodology used to collect and analyze the state and local data presented in this report. Section II is an overview of the assessment and accountability systems in New York, including changes resulting from NCLB. In Section III, we provide demographic information on each of the study sites. In Section IV, we present our research findings at the state and district levels, and in Section V we present a discussion of the major findings for each embedded case study and address cross-site themes.

I. METHODOLOGY

Identification of Study Sites and Data Sources

The Education Policy Reform Research Institute (EPRRI), funded by the United States Department of Education's Office of Special Education Programs (OSEP), was established in 2000 to investigate and describe the impact of including students with disabilities in updated state educational accountability systems as required by the 1997 reauthorization of the Individuals with Disabilities Education Act (IDEA, 1997). EPRRI researchers adopted the case study method as an empirical approach to answering the research problem and specified a priori the levels of analysis, participants, and questions to be pursued. One prerequisite for selection was the requirement that study states in the study have high-quality data collection systems capable of providing disaggregated data for students with disabilities.

Four core study states agreed to participate in the five-year study, these states were California, Maryland, New York, and Texas. Within each state, EPRRI staff and core state representatives identified and secured the participation of two school districts in the study. The four selected states varied across several key accountability features, including high-stakes versus low-stakes accountability consequences, recentness of reforms, stability versus instability of reform efforts, participation of students with disabilities in all accountability reports, and use of alternate assessments. In addition, the sites also reflected geographic and demographic diversity. It is important to recognize that in December 2001, the policy context surrounding this study changed substantially with the passage of NCLB and as a result the study sites became more homogeneous in their key accountability features. Unfortunately, due to our initial study design we had

already selected sites and participants on an a priori basis (versus selecting them purposively and serially over time) and it is possible that we limited our opportunities to study relevant issues that only became apparent after NCLB and over the course of the study. Despite these reservations, we are confident that our analysis of data obtained from selected states, districts, and schools reveals important issues and “key” challenges as states, districts, and schools nationwide work to meet the ongoing requirements of NCLB.

EPRRI researchers utilized two complementary strategies to collect qualitative data: analysis of documents and in-depth interviewing (Marshall & Rossman, 1999). The use of two methods of data collection and multiple data sources allowed for triangulation by data source. EPRRI staff developed a sound understanding of the history and context of educational reform in each of the study sites by ongoing reviews of state and district policies available online or obtained directly from the participants. Examples of the types of state documents reviewed include Board of Education policies and minutes, strategic plans, reports from Superintendents and Commissioners, and state Education Department letters to districts. We also reviewed state-consolidated applications for federal funding under NCLB and state accountability plans submitted to the U.S. Department of Education to comply with NCLB. In addition, we collected press releases and letters issued by the Secretary of Education, and speeches by the President and Secretary that were obtained from the U.S. Department of Education web site. We also reviewed reports issued by the U.S. General Accounting Office on NCLB, journal articles and local newspaper articles from across the four sites. Finally, we monitored state Web sites for

information on evolving state efforts to implement NCLB and for state policy documents developed to meet the NCLB requirements.

The guiding questions for the research came directly from the Office of Special Education's grant priority, but were further broken into subheadings and sample issues/indicators. In February 2001, the state-level participants and EPRRI staff together reviewed and revised a draft template of research questions. This is available on the EPRRI website, www.eprri.org. This process led to the development of 10 individual interview protocols focused on the areas of accountability, assessment, monitoring, curriculum and instruction, special education, teacher certification, professional development, Title I, transition, and parent viewpoints. In October 2001, the same process was used to develop interview protocols at the district level. Individual school level protocols were developed for the school principal, regular education teachers and special education teachers in late 2002.

Selection of participants at the state and district level was done purposively with the support and cooperation of the state and district directors of special education, who acted as gatekeepers. At the state and district levels, the special education director participated in the identification, initial contact, and interview arrangements with key personnel.

Knowledgeable personnel from the following departments in each SEA and LEA were interviewed: special education; accountability; testing; special education monitoring; Title I monitoring; curriculum; teacher certification; and professional development.

Overview of Data Collection Procedures

During this five-year multi-state study, we conducted 35 interviews at the state level and 44 at the district level. State level interviews occurred between October 2001

and January 2002 and district level interviews took place between April and June 2002. The interviews at all levels of the education system were between one and two hours in length and adopted a semi-structured approach with the interview protocols acting as rough guides. Interviews were also audiotaped with the participants' consent so that they could take the form of a conversation. Finally, the interviews were conducted by multiple researchers (Eisenhardt & Bourgeois, 1988), allowing the researchers to overlap data analysis and data collection, as recommended by Glaser and Strauss (1967) as well as Van Maanen (1988). The use of multiple investigators provided complementary insights, added richness to the data analysis, and enhanced confidence in the findings (Eisenhardt, 2002).

Data Analysis

EPRRI researchers followed the qualitative data analysis procedures recommended by Miles and Huberman (1994). After each interview, field notes were written-up by individual researchers, while the taped interviews were transcribed by graduate students. At each level of the system individual case study, data analysis began with the creation of a contact summary sheet to develop an overall picture of the main points of each interview. The graduate assistant who first transcribed the interview then read each contact summary sheet to identify bias and selectivity. We developed detailed descriptive write-ups for each site, based on the field notes, the contact summary sheet, and the document reviews. This step in the data analysis is central to the generation of insight and helps researchers to cope with the enormous volume of data (Miles & Huberman, 1994; Pettigrew, 1988).

We then entered the transcripts into a qualitative software program called

Ethnograph, which allows for the analysis of text-based data into codes and categories of meaning. EPRRI researchers adopted a coding approach partway between the a-priori and inductive approaches discussed by Miles and Huberman (1994). A general coding scheme, based on that developed by Bogdon and Biklen (1992), was created to provide a structural, conceptual, and coherent order to the emerging codes. This particular coding scheme was based on the general domains covered by the research matrices and the interview protocols and allowed codes to develop inductively, while at the same time enabling the researcher to “think about categories in which codes will have to be developed” (Miles & Huberman, p. 61).

We created clear operational definitions for each code so that the codes could be applied consistently. Code names closest to the concepts being described were applied to the chunks of data. Initial coding of the data corpus was performed by the first author and a team of graduate research assistants, who read and reread each interview line by line and coded the sentences or phrases relating to the participants’ perceptions of the effects of accountability on students with disabilities and the systems that serve them. EPRRI researchers followed the recommendations of Miles and Huberman (1994) for check-coding. Two researchers separately coded the first eight pages of the opening interviews from each level of the system and reviewed the coded sections together. Intercoder reliability was determined using the following formula:

Reliability = $\frac{\text{number of agreements}}{\text{total number of agreements} + \text{disagreements}}$.

Initial intercoder reliability at each level ranged from 79 to 83% and rose up to between 91 and 93% after the differences were clarified. A further check was performed two thirds of the way through the data analysis. A conceptually clustered matrix was

developed during the early analysis based on the interview protocols. The following decision rules were applied as themes were identified: first a theme was coded as present for a participant if it was mentioned repeatedly or with strong emphasis during the interview and second, a theme was coded as present for a study site if it was mentioned by 2 or more participants.

The second step in the analysis was to put the data back together again in a new way to reveal themes and stressors related to the impact of accountability at each level. This process is similar to axial coding in grounded theory (Strauss & Corbin, 1990), involved grouping and comparing the initial codes with each other and merging similar concepts together into larger encompassing themes. During this process, all key ideas, findings, and interpretations were presented and discussed by the EPRRI staff and, at the state and district level, with study participants (Miles & Huberman, 1994). Original transcripts were used as evidence to support the emerging themes, and at all times the researchers searched for examples that would contradict key findings (Yin, 1989).

II. NEW YORK ASSESSMENT AND ACCOUNTABILITY SYSTEMS

Each of the four states that participated in this research has a unique accountability system to determine the academic achievement of their students. Information about each state's assessment system was retrieved from policy documents and other sources, primarily the Internet. Examples of the types of documents reviewed include Board of Education policies and minutes, strategic plans, reports from Superintendents and Commissioners, Department of Education memos, policies relating to standards-based reform, assessments, and accountability, as well as press releases and

other reports. Following is a description of the assessments and accountability systems in place in New York State up to the end of the 2003-04 school year.

The Board of Regents and its State Education Department (SED) govern education from pre-kindergarten to graduate school. Together they are constitutionally responsible for setting educational policy, standards, and rules—and are legally required to ensure that the entities they oversee carry them out. The Board and Department also govern 38 licensed professions; provide vocational and educational services to people with disabilities; guide local government records programs; oversee public broadcasting; and operate the State Archives, Library, and Museum.

The state of New York has a history of school improvement initiatives beginning with the 1984 *Action Plan to Improve Elementary and Secondary Education Results in New York* and *A New Compact for Learning*. However, in New York the state provided a framework for the local school districts—leaving curriculum and instruction to the district, but having quality control through assessments at the high school level. In 1996, the Board of Regents developed a new strategic plan for public education entitled *Leadership and Learning*. Standards-based reform elements were at the center of the Regents strategy including the need to: (a) set clear, high expectations/standards for all students and develop an effective means of assessing student progress in meeting the standards; (b) build the capacity of schools and districts to enable all students to meet standards; and (c) use and expand the existing systems of public accountability for schools, based on student performance, and provide incentives for improving effectiveness and sanctions for low performance.

The New York Learning Standards

Until 1996, local districts had complete discretion over curriculum and instruction, although they needed to be cognizant of the content area requirements at the commencement level for students to be successful on the Regents exams or the Regents Competency Tests (RCTs). As part of its standards-based reform plan, the Board of Regents approved and adopted a set of 28 learning standards for New York students that formed the basis of their revision of education in New York. These standards represented the core of what all people should know, understand, and be able to do as a result of their schooling. The goals of this initiative were to provide a rigorous and intellectually challenging education statewide; closely align teaching, assessment, and the provision of supports for learning; and create an environment where schools and parents shared the same high expectations of youngsters. This marked a shift in education policy, which had until then largely been determined at the local level.

The 28 standards focused on seven curriculum areas: English language arts; The Arts; Social Studies; Math, Science, and Technology; Languages other than English; Health, Physical Education, and Family Consumer Sciences; and Career Development and Occupational Studies. These learning standards formed the basis for the education of all in New York (www.nysatl.nysed.gov/standards.html). All children were expected to acquire a working knowledge of each area and develop proficiency in applying this knowledge to meaningful tasks. Following the adoption of the standards, the SED issued a series of core curricula for each learning standard to present key ideas (broad, unifying, general statements of what students need to know) and performance indicators (statements of what students should do to provide evidence that they understand the key

ideas). These core curricula are the foundation upon which state assessments are aligned and developed. New York State core curricula were not designed as local school/district curricula. Rather, these core curricula provided a framework to local schools/districts, who maintain responsibility to design curricula that meets the needs of their students. The core curricula respect the tradition of local choice in New York State, within which educators select texts, identify products, and use a rich array of instructional strategies and activities to meet student needs.

New York State Assessments

As part of its education reform initiative, the Board of Regents raised performance expectations at all grade levels and redefined the requirements for high school graduation to align with these new performance expectations. New York's policy attached high stakes for students through the Regents high school exit exam (DeBray, Parson, & Woodworth, 2001). Assessments are administered at 4th grade and 8th grade in English Language Arts (ELA) and Mathematics; this began in May 1999. These new tests replaced the Pupil Evaluation Program tests in Mathematics, Reading, and Writing in grades 3, 5, and 6 and the Preliminary Competency Tests in English given previously in grades 8 and 9. Students are also tested on Social Studies in 5th grade and 8th grade, Technology Education in 8th grade, and Science in 4th grade and 8th grade.

The assessments were developed by committees of teachers, testing specialists, and subject specialists. The test questions are in a variety of formats including multiple choice, short answer, and essays. The test questions were reviewed by a special committee to ensure that they were free from bias and were culturally sensitive. The assessments are scored regionally by teachers from New York State. In response to

NCLB, annual testing of students in grades 3-8 was introduced in 2005-06 in ELA and mathematics.

Initially, SED allowed a range of accommodations on the state assessments for students with disabilities. However, SED issued a memo in February 2002 informing districts about changes in allowable testing accommodations on elementary and intermediate level tests. Beginning with the 2002-03 school year, read-aloud was no longer permitted for use on sections of the state's 4th grade and 8th grade ELA tests that measure reading comprehension. In addition, students with disabilities were not allowed to use a calculator on the grade 4 Mathematics assessment or on the first part of the grade 8 Mathematics assessment. The state revised its accommodations policy because use of these specific accommodations on certain sections of the assessments changed the construct of the test and yielded invalid scores. One of the potential consequences was that students with disabilities might not receive needed supports and access to Academic Intervention Services. In addition, these accommodations yielded scores that did not provide teachers with valid information to guide instruction.

Until 1996, high school students earned either a "local" or "Regents" diploma. The local diploma was awarded after passing five Regents Competency Tests (RCTs); a Regents diploma was awarded if students passed eight higher-level subject matter exams. As part of its overall plan to improve education, the Board of Regents announced that the RCTs would be phased out over a period of eight years and all students would be required to pass five higher-level Regents exams in English, Mathematics, Global Studies, U.S. History and Science with a score of 65 or higher, in order to graduate. A score of 55 to 64 can also be considered a passing score for a local diploma up to the

2007-08 school year. This decision is made by individual school districts. A safety net policy is in place for students with disabilities who first enter 9th grade in or after September 1996 and prior to September 2010. This safety net allows students with disabilities who take and fail the Regents examination in a required area to meet the requirements for a local diploma by passing the Regents Competency Test in that content area with a score of 55 to 64.

The revised Regents examinations were designed to prepare students to begin college-level study or skilled employment. They require more conceptual understanding, problem solving, application of knowledge, and critical analysis than was previously required. Because of the greater rigor of the revised assessment and the high stakes for students, New York allows component retesting for any senior student who fails a Regents examination twice and has scored at least 48 on an examination. Schools are able to analyze the results of a student's tests and determine the specific areas of the standards where he or she needs additional instruction. There are two component tests in ELA and four component tests in Mathematics. Students are provided with remedial help and then retake the components that they did not pass on the full Regents examination.

Alternate assessment. In 1999, the State Education Department began development of an alternate assessment for students with severe disabilities, as required by IDEA 97. A statewide task force of educators, researchers, parents, and advocates, along with the State's alternate assessment testing contractor, developed guiding principles, process and participation criteria, guidelines for creating assessment tasks to measure progress on the alternate performance indicators, and a scoring rubric. Students with severe disabilities may be declared exempt from the general assessment by the

Superintendent and Committee of Special Education (CSE) based on the following criteria: (a) severe cognitive disability, significant deficit in communication/language, or significant deficits in adaptive behavior; (b) requires a highly specialized educational program that facilitates the acquisition, application, and transfer of skills across natural environments; and (c) requires educational support systems, such as assistive technology, personal care services, health/medical services, or behavioral intervention. These decisions are made on a case-by-case basis.

The New York State Alternate Assessment (NYSAA) was created for students with severe disabilities who are at the ages of 10-11, 14-15, or 17-18. The NYSAA is a data folio assessment in which students demonstrate their performance toward meeting the alternate performance indicator level of the New York State Learning Standards. Like the alternate assessment in many states, the NYSAA was designed for only a small group of students with disabilities and was developed in response to IDEA 97 legislative mandates. However, to comply with NCLB, SED must make changes to the NYSAA, to be implemented during 2005-06 school year, that include requiring students with severe cognitive disabilities to be assessed according to birth dates that are aligned with grade level equivalents from grade 3 through grade 8 (NYS Memo, September 2005). In New York as in many states, policies relating to alternate assessments remain very fluid as the state attempts to comply both with NCLB and with the recently reauthorized Individuals with Disabilities Education Improvement Act of 2004 (IDEA).

In March 2002, the New York State Department of Education added a provision that permitted locally selected assessments to be used in lieu of the state assessments at 4th and 8th grade and for commencement-level assessments. This option was available

only to certain students with disabilities who, due to a performance gap resulting from the students' unique disability needs, were unable to meet the grade/age level expectations of the regular state assessments and did not meet the eligibility criteria for the NYSAA. The locally selected assessments must be standardized, measure a student's achievement of the New York State learning standards, and be of sufficient technical quality to measure progress toward the student's performance indicator level. For purposes of accountability, students who participate in locally selected assessments are counted as though they performed at level one on the New York State Assessment Program (NYSAP) elementary and middle school ELA and mathematics assessments. In its comments on NCLB proposed regulations for alternate assessments, SED stated that it believed that the percentage of students with disabilities who would need an alternate assessment was larger than that proposed by the U.S. Department of Education. SED reported that the figure was between 20% and 30%.

Locally selected assessments were first available for an initial two year period for "certain" students with disabilities in grade 4, grade 5, grade 8, and at commencement level, while the SED conducted pilot elementary and intermediate assessments of these students who have been recommended by their Committees on Special Education (CSEs). If a student participated in a locally selected assessment in lieu of a Regents or an RCT, then he or she would not meet the assessment requirements for a local or a Regents Diploma and would be eligible for an IEP diploma. In April 2003 the provision was extended for the 2003-04 school year; in June 2004 it was extended to include the 2004-05 school year; and was since extended through the 2009-10 school year.

For the 2005-06 school year only, SED issued interim guidelines necessary to meet accountability requirements under NCLB pending further action from the U.S. Department of Education regarding modified achievement standards. This interim guidance identified a group of students with disabilities who “may need more time before their participation in the general state assessments and who do not meet the participation criteria for NYSAA.” Under this interim measure the CSE may determine for state assessment purposes only that a student with a disability is graded or ungraded. To be considered ungraded a student who is ineligible for NYSAA must have an instructional level “significantly below (three or more years) the grade-level coursework in ELA and mathematics.” If a student is determined to be ungraded for the purposes of participation in the state assessments, participation is based on his or her date of birth. To illustrate, a student born between July 1, 1992 and June 30, 1993 would take 5th grade ELA, Mathematics, and Social Studies. The U.S. Department of Education’s recognition of another group of students with disabilities accorded well with SED’s position on the December 2003 Final Rule. In its comments on the regulations, SED argued that the 1% cap was too low and that the New York State data showed that there was a substantially larger group of students with disabilities who would not reach grade level expectations (www.regents.nysed.gov/2004Meetings/November2004/1104brd5.htm)

New York State Accountability System

The Board of Regents became concerned about the standard of education in the mid-1980s and developed an action plan requiring more courses of study and more credit hours to graduate with a Regents diploma. In the 1990s, the Regents begin to define at the state level what the outcomes of the educational system should be, while allowing the

local school districts to have discretion on curriculum and instruction. The Regents adopted standards-based reform in the late 1990s with a focus on what students need to know and be able to do, and developed Statewide Learning Standards and graduation requirements. When the Regents first adopted standards-based reform there was a strong funding response—a more than \$2.6 billion increase in state aid to school districts from the state legislature.

In 1999-2000, a System of Accountability for Student Success (SASS) was established. This system became New York State's single accountability system, designed to provide information about school effectiveness in preparing students to meet New York State's learning standards, as indicated above. These standards are aligned with curriculum and the NYSAP across elementary, middle, and high school levels. At the high school level, schools are held accountable for (a) the percentage of students who graduate with a Regents diploma in four years and (b) decreasing the annual high school dropout rate.

New York established four levels of performance on the NYSAP to determine a school's Performance Index, which reflects achievement toward state standards. Students receiving a score of Level 1 are identified as having serious academic deficiencies and as having shown little or no proficiency in the New York state content standards for their grade level. Level 2 designates students who have shown some knowledge and skill in each of the required state standards for elementary or middle school level students, but who still need extra assistance to achieve all of the standards required to pass the assessments. A Level 3 designation means that a student has met standards, while Level 4 means that a student has exceeded standards. The scores of students who participate in

the NYSAA are included, with a maximum of 1% of the proficient scores used in calculating the Performance Index.

The performance of all students, regardless of whether they take the standard or alternate assessment, must be reported through the Local Educational Agency Program (LEAP). LEAP is an electronic software tool that collects and reports on State assessments administered in the elementary and middle school levels. By using this tool, a determination is made of the school's Performance Index, which is the percentage of full-year students tested who scored at Level 2 and above, and the percentage who scored at Level 3 and above on each assessment. This determines whether the school has achieved the state standard and its AYP target. Another data collecting and reporting tool is the System for Tracking Educational Performance (STEP). STEP is software designed to track all students' performance in grades 9-12. This tracking tool also determines if schools have met the state standard of the Performance Index and AYP.

Prior to NCLB, AYP in New York State was determined by using the performance of Title I schools rather than by measuring the performance of all students. A district made AYP criterion if at least 50% of its schools met the state standard or reached AYP on an assessment. Since the No Child Left Behind Act was established, changes have been made to STEP. The New York State English as a Second Language Achievement Test (NYSESLAT), NYSAA, and the graduation/dropout rate are now included in the system's accountability measures and the performance index is increased. Now the district is held accountable for all student subgroups, including individuals who have been placed in programs outside of the district (i.e. approved private placements). Also, prior to NCLB, there was a single measure for elementary and intermediate level

ELA and a single measure for elementary and intermediate level Mathematics that combined results of grade 4 and grade 8 assessments. Under NCLB, these measures are now separated by grade level.

LEAP and STEP determine if a school has met each state standard and then labels each school as ‘Meeting Standard’, ‘Below Standard’, or ‘Farthest from Standard.’

Schools that are farthest below state standards are designated as Schools Under Registration Review (SURR). SURR, which was developed before NCLB, is a program created by the Board of Regents to close the gap in student performance. A SURR school is labeled “School in Need of Improvement” if the school performs below state standards and fails to make AYP for two consecutive years. A school may also be labeled a “Corrective Action School” if the school, which was previously under “School in Need of Improvement” designation, fails to make AYP in two out of the next three years.

SURR was created prior to NCLB, as noted above, and the program has changed since the implementation of NCLB. Schools used to be expected to demonstrate that 90% of their students were achieving at the minimally acceptable performance level (Level 2). If a school did not demonstrate this, the school was placed on registration review and given three years to improve performance. If no improvement was evident, the school had to close or reorganize. However, these schools were not provided with substantive guidance on how to reorganize. Now, a Local Assistance Plan (LAP) and AYP target must be established for the school to raise its level of compliance with state standards. The district in which the school is located develops the LAP. Since NCLB, there is a need for additional assistance to ensure that a sufficient percentage of the subgroups within a SURR school will achieve proficiency in ELA and Mathematics assessments. In addition,

the Commissioner established the AYP target for each school below state standards. For elementary and middle schools that are below the targeted performance level, the targets are set in three-year increments and the school is expected to close the gap by 15% each year. For high schools, the goal is to demonstrate improvement over the previous year's performance.

III. NEW YORK STATE AND LOCAL DISTRICT LEVEL DEMOGRAPHICS

In the 2000-01 school year, New York State enrolled just over 2.8 million students and identified just over 38% as economically disadvantaged (Table 1). Approximately 12% of New York State students received special education services and 8.4% services for English learners. About 55% of New York's students were White, just over 20% were African American, 18% were Hispanic and the remaining were Asian, American Indian, or other. The average student-to-teacher ratio was 13 to 1, and the average operating expenditure per pupil was \$6,150 (Table 3).

Enrollment in New York State remained fairly stable in the next four years and totaled nearly 2.9 million students in the 2004-05 school year, with nearly 37% of these students being identified as economically disadvantaged (see Table 2). The percent of students receiving special education services increased slightly to 14%, while the percent of students receiving bilingual services decreased to just over 7%. The diversity of the student body in New York State also remained relatively stable over the years. By the 2004-05 school year, about 53% of students were White, 20% were African American, an additional 20% were Hispanic, and the remaining 7% were Asian, Pacific Islander, or American Indian. The student-to-teacher ratio continued to be around 13 students per

teacher, although the operating expenditures increased to just over \$8,000 per student.

Table 1: Selected state and district demographics, 2000-01 school year

State/ School District	Enrollment	% FARMS	% Students Receiving Special Education Services	% Students Receiving English Learner Services	Ethnic Group			
					% African American	% Hispan ic	% White	% Other
New York	2,823,602	38.1	11.8	8.4	20.1	18.4	55.1	6.4
North Colonie	5,500	6.3	9.1	2.5	3.2	1.4	88.5	6.9
Rochester City	37,885	67.2	15.1	6.5	62.8	18.9	16.1	2.2

FARMS=Free and Reduced Meals

Table 2: Selected state and district demographics, 2004-05 school year

State/ School District	Enrollment	% FARMS	% Students Receiving Special Education Services	% Students Receiving English Learner Services	Ethnic Group			
					% African American	% Hispan ic	% White	% Other
New York	2,880,411	36.6	14.0	7.2	19.9	19.7	53.2	7.2
North Colonie	5,585	6.0	10.5	1.9	4.1	1.7	87.0	7.3
Rochester City	33,055	68.6	18.8	7.5	64.9	20.1	13.0	2.0

Table 3. Selected state and school district staff and financial statistics, 2000-01 school year.

State/School District	Average Number of Students/ Teacher	Total Expenditure per pupil for general education
<i>New York</i>	<i>12.9</i>	<i>6,150</i>
North Colonie Central	14.3	5,425
Rochester City	11.7	5,547

Table 4. Selected state and school district staff and financial statistics, 2004-05 school year.

State/School District	Average Number of Students/ Teacher	Total Expenditure per pupil for general education *
<i>New York</i>	<i>13.1</i>	<i>8,177</i>
North Colonie Central	13.6	6,636
Rochester City	10.9	6,894

* Data from 2003-04 school year

North Colonie Central School District is a small suburban district located just outside Albany. In 2000-01, about 6.3% of students were economically disadvantaged, a much lower percentage than that of the state of New York overall. About 9% of North Colonie students received special education services and a very small percentage (2.5%) received English learner services. Almost 90% of the students were White (88.5%). The number of students per teacher (14.3) was slightly higher than the state overall average, and the operating expenditures per pupil (\$5,425) were slightly lower than the state average.

By the 2004-05 school year, the enrollment in North Colonie had increased slightly, while the percent of students identified as economically disadvantaged decreased by a small margin to 6.0 %. About 10% of students received special education services, and about 2% received services for English learners. The county continued to enroll a high percentage of White students (87%) in the 2004-05 school year. The student-to-

teacher ratio continued to be higher than the state average, although it decreased slightly to 13.6 students per teacher. The operating expenditures per pupil for the 2003-04 school year increased to \$6,636, but still fell below the state average.

Rochester City School District is located in Monroe County and is one of the Big Five schools districts in the state, along with New York City, Buffalo, Syracuse, and Yonkers. Over two-thirds of the students in Rochester were economically disadvantaged in 2000-01, a much higher proportion than the state as a whole. Fifteen percent of the students in Rochester received special education services and 6.5% received English services. Almost 63% of the district's students were African American, 19% were Hispanic, and White students represented 16% of the total student body. The student-to-teacher ratio was 12 to 1, which was close to the state average, with an operating expenditure per pupil of \$5,547, lower than the average for the state as a whole.

By the 2004-05 school year, enrollment in Rochester County had decreased to just over 33,000 students. A high percentage of these students continued to be identified as economically disadvantaged (68.6%). The percent of students receiving special education services and those receiving services for English learners both increased slightly, to 18.8% and 7.5%, respectively. The percent of African American students in Rochester increased to about 65% in the 2004-05 school year, while the percent of Hispanic students increased to just over 20%. In turn, the percent of White students decreased to 13%. The student-to-teacher ratio also decreased to about 11 students per teacher, which continued to be lower than the state average. The operating expenditures per pupil increased to nearly \$7,000, but remained lower than the state average.

IV. NEW YORK STATE EMBEDDED STATE AND DISTRICT FINDINGS

Higher performance standards for all students, especially at the high school level, and decreasing the statewide performance gap, were at the heart of New York's educational reform plan and reflected the basic tenets of standards-based reform, including the belief that performance information from the accountability system would drive change in schools and districts. The data from New York State indicate support for accountability in theory, but also indicate the presence of a key limitation to the practical application of this theory: the lack of existing local capacity to implement change. Findings in this section are arranged into three broad topical areas: (a) making accountability work for students with disabilities, (b) building instructional capacity, and (c) setting state and district responsibilities. Within each theme are related subthemes that vary according to the time period and level from which data were collected.

State level interviews took place in October of 2001 and district interviews took place in May of 2002. EPRRI staff interviewed eight individuals at the state level; in North Colonie we interviewed 10 district personnel, and in Rochester we interviewed 11 district personnel. In each topical area, the data will generally be presented at the state level and then the district level. The significant conceptual themes that emerged will be illustrated by a small number of exemplar quotes. Each quote was chosen to be representative of many such instances, or chosen because it was an extreme instance or the sole instance or represented two contrasting properties of the conceptual theme under discussion.

Making Accountability Work for Students with Disabilities

New York State reported the performance scores for students with disabilities on

state assessments before being required to do so by NCLB. The state has reported these scores since the 1998-99 school year. When the state level and district level interviews took place in October 2001 and May 2002, respectively, our informants were concerned with making accountability work for students with disabilities in the same way it was supposed to work for students without disabilities. In the following section, we will explore three themes: (a) increasing educator expectations for this population, (b) providing information to stakeholders, and (c) using data to make instructional decisions. Informants from each level addressed the elements of accountability reform that were more problematic for this population, such as high-stakes graduation examinations, accommodations on state assessments in grades 4 and 8, and meeting the needs of students with disabilities who fell between the state's regular assessment and the state's alternate assessment. Informants labeled this group the "grey kids."

Increasing expectation. One of the important premises of standards-based reform is the requirement that schools hold high expectations for all students. However, informants from the state and district agreed that expectations for students with disabilities had always been low. The following comment from a state level informant is illustrative of many we heard at both levels of the system and reveals the cyclical nature of low expectations leading to instructional programs that lack educational rigor, leading in turn to poor performance and continued low expectations. One of the promises of accountability reform is the hope that this cycle can be broken and children with disabilities can have the opportunity to demonstrate their abilities:

So I think, historically, that we underestimate the ability of children with disabilities generally. We usually set up our instructional programs around those estimates and expectations, which mean that they are almost always lower than

they should be. We want to force people to reconsider their expectations and then get on with allowing children to have opportunities they haven't had before.

At the district level, visited a year later, informants from our two study sites reported that most students with disabilities were an integral part of standards-based reform and that the superintendent in each district had made this clear from the beginning: "What is interesting is our superintendent has been very firm, and when he talks about kids he's talking about all kids. That's why you won't see a whole lot of 'and kids with disabilities.'" Because of this inclusive approach, most students with disabilities had access to the general education curriculum:

Our approach wasn't that kind of segmented approach, it was a general approach toward high standards. With standards-based instruction it's a high standard for everybody. It's not a special kind of material for a special kind of folk. There's no special curriculum.

According to our district informants, only students with severe disabilities were given access to a different curriculum:

There is a small portion of our student population in special education that have a more specialized curriculum that's more appropriate for activities of daily living, but it is still aligned with the standards. But other than that we don't have any specialized curriculum.

As the following statement reveals, the commitment to including students with disabilities in accountability reform was made by the district's superintendent. This individual instilled the belief that the issue was not if this population was included, but how to achieve this goal:

To give our superintendent credit, I think he really came on very strong in terms of making sure that all students have the same expectations, including students with disabilities. That was right from the beginning and that made a big difference because that stage was set and so there wasn't a debate over whether or not, just how are we going to?

There was evidence from both districts that children with disabilities were slowly beginning to meet the challenge of higher standards. This trend was clearer in the smaller, more affluent district, but was also apparent in the larger urban district, as this response illustrates:

Many students with disabilities are embracing the idea that they have the opportunity to be in Regents courses. That they're thought of along with everybody else and that no, they're not getting an IEP diploma. They're saying: "We want a real diploma." That's a wonderful sense of self-awareness and self-advocacy that hasn't been there in the past.

However, as a different informant pointed out, there was still a long way to go in increasing expectations for this population. For example, very few students with disabilities were recommended for Advanced Placement classes or International Baccalaureate classes. According to this informant the problem lay with general education:

We have some children with disabilities in our AP courses and our international baccalaureate programs, but very, very few. And I'll be real blunt here, I think the reason is that guidance counselors do not think of them in those terms, unless they truly are the exceptional. The battle that we really have been fighting in this district is in the general education curriculum, and we're still fighting that now. We haven't necessarily won that one.

We asked our informants if the performance of students with disabilities was improving because of the state's educational reforms. In our more affluent district informants were very positive, but attributed the high performance of students in special education to the overall culture of the district and not the state's initiatives. In our urban school district informants reported that there was an upward trend in the data for students with disabilities, but that there was still a long way to go for all students. Interestingly, as the following quote illustrates, some students in special education outperformed their peers without disabilities:

The trend is an upward trend. We still have way too many kids who are not meeting the standards, but one of the things we are seeing that we didn't have is movement towards meeting the standard. In fact, when we look at the assessments, the students with disabilities are actually doing better than many of the general ed kids. It's taking time to change people's thinking, and to demonstrate to teachers that students with disabilities, when given the different strategies, the right kinds of accommodations, differentiated instruction and so forth, can perform and can reach standards.

Providing information to stakeholders. Another important component of standards-based reform is transparency of results to enable stakeholders such as parents to evaluate their local system and to make informed educational decisions with or for their children. Public reporting assumes that if local stakeholders are uncomfortable enough with low performance, they will provide the necessary local-level pressure to bring about improvement. Findings from this study indicate that parents of students with disabilities are asking questions about the performance of this population in the school system and expecting that school systems will address their needs. The following comment is an excellent example of how parents of students with disabilities can use

publicly available school report cards to hold school systems publicly accountable. These school report cards are a source of information previously unavailable to parents of students in special education. According to this informant, making information available to stakeholder groups provides an impetus for change:

When the assessment scores come out, every single newspaper in the city will publish them district by district, including special education scores. I've had parents who know me come to me and say: "Did you see how that school's special education student graduation rate is twice as high as mine?" I say well that's interesting, what do you think? And they say: "Well I'm going to go to the next school board meeting and ask them." People are asking, what are you going to do to improve? And then the next day the superintendent announces a new initiative.

Another state level informant confirmed the impact that public reporting of performance scores had on school systems. Moreover, this informant, originally opposed to public reporting, commented that it was a necessary component of educational reform: "I was against public reporting, but I really believe it is the wake-up call. Once those special education scores are posted, we get the calls. It isn't until those scores are made public that people say, here's where we need to focus."

The importance of public reporting also emerged at the district level. However, as this informant indicated, people were initially shocked at how poorly students with disabilities were doing in the district:

People didn't know before how badly kids with disabilities were doing and now they do and it's not really acceptable to a lot of people. Our superintendent was very unhappy with this stuff, but he didn't even know about it until recently. You are going to see really measured efforts to improve those scores. More and more

people are less and less satisfied with poor results.

When we visited the district in May 2002, we asked informants for their views on the recently implemented NCLB. Although the federal regulations accompanying the statute were not available, informants were just beginning to get a sense of the impact of this major piece of federal legislation. As the following comment indicates, our more affluent district's informants were not overly concerned about NCLB, as this district was performing very well on the state assessments:

We haven't changed anything because of NCLB because I think we had everything in place. So, it really didn't make a huge impact on us, honestly. We're doing what we can, and I think what we're doing is right. At this point, we're doing a good job.

In our other district, NCLB received mixed reviews, although most were in favor of the legislation in theory. As one informant pointed out, although reaching 100% proficiency was a challenging goal, the target had to be 100% or a group of students would immediately be left behind:

I think if we don't say that it's doable then it becomes impossible. So that what you've got to do is make it as doable as you can and get as close as you can. You can't a priori establish like a 70% criterion and say that's all that can make it because then you're excluding a population.

On the other hand, a different informant from this district believed that the comparable improvement component of AYP would be very difficult to achieve. In addition, this informant believed that people would try to exclude low performing students to make their schools look better:

I see a whole lot of people not knowing what the hell to do. You're going to find folks who are going to be jumping up and down trying to figure out which kids not to include and how to make their performance look better. You know you're going to get that.

Using data to make instructional decisions

A third component of standards-based reform is the use of performance information to drive instructional change at all levels of the educational system. As the following quotation illustrates, state level informants supported this approach and were hopeful that it would succeed: "We are hoping that one of the results of all of this data, and obviously much more clearer and informative data now than we have ever had, is going to push people to improve their instructional programs." Another participant commented that she was very optimistic about the shift from procedural compliance to results, pointing out that outcomes for students with disabilities were abysmal in many instances:

I'm very optimistic about what's going on, as we have got away from the procedural compliance to looking at improving outcomes. Our new benchmarks and triggers have to do with performance, percent of kids who take the assessments, kids who graduate, dropout rates, all that kind of thing. If you look at the results that we've had for our children as adults it's abysmal. So who cares if we did their IEP on time?

Finally, a participant at the state level emphasized that the state was focusing in on key results and applying pressure to districts to improve performance. This informant also reflected that public reporting made it much harder for districts to claim that it was impossible to improve the performance of students in special education, because evidence from other districts proved them wrong.

We are really focused on certain results and we are coming at the districts from every angle we have, always on the same thing. That to me is what has really begun to make a difference. People know things about special education that they didn't know before. We really published this information incessantly, and at first everybody got very angry and then it kind of wore everybody down and now what people are doing is responding. We have tried very hard to publicize what and who's been effective because the more we do that, the less willing people are to stand up and say this isn't fair and we can't do this.

When we visited the districts a year later we asked informants whether schools looked at performance data from state assessments to inform instructional practices for students with disabilities. Responses from our two districts indicated a difference in the level of data use between the two study sites. The following two quotations illustrate a difference in the comfort level regarding using data to guide instructional decision-making. In the first quotation our informant emphasized that the district looked at and used performance results on the state assessments to inform individual student and school programming and to guide staff development:

It's very important to us to take a look at the results of our state assessments. As soon as those come in, we take a look at them, not only looking at individual students and classes, but also the entire grade level. We do some analysis district-wide, and if we need to provide staff development for some of our teachers, we then add that piece into some of our staff development days.

As the second quotation illustrates, our other district was moving toward data-driven decision-making, but was not yet at the stage where its use was institutionalized at all levels of the system.

There's been a movement towards looking at student performance data at various levels of strata and particularly at gaps in performance. We are just one layer and then another layer of leadership training where principals and other administrators are getting more information and training with data-driven instruction.

Eventually, all levels of the organization will be focusing on student data. What does that tell us about our pedagogy, how do we change our practice for better results? There is a problem in education in general in moving into the world that we are describing. We're really not clued into real accountability of looking at gaps and trying to improve them. We're into that world but we're not totally comfortable with it yet.

High stakes graduation and students in special education

The new tests with high stakes for students were perhaps the most prominent feature of the state's standards-based reform initiative. Along with multiple opportunities to take the exit level Regents and the availability of component retesting, state level informants described how students in special education were protected from the consequences of increased graduation requirements through the safety net. Unlike their peers without disabilities, students in special education could take an RCT after they had failed a Regents examination in that subject. If a student scored between 55 and 64 on the RCT, he or she could graduate with a local diploma. Students without disabilities could graduate with a local diploma if they scored between 55 and 65 on the Regents examinations. As the following comment illustrates, policy makers were concerned about the negative impact on students with disabilities of the state's commitment to high-stakes graduation requirements. According to this informant, the state policy makers needed to decide whether to fully include this population in standards-based reform, which would mean attaching high-stakes consequences to them or making it possible for all students in

special education to graduate with a local diploma:

The question then is a policy question. Do we want every kid in special education to have a way to get a local diploma, and therefore in effect to go back to lowering expectations and state standards. That will be a very big issue, but I think is still down the road a question for us because we still have our safety net in place.

At the time of our district visit, the state was considering extending the safety net provision, a move that both our districts fully supported: “Well, we all hope that the state will agree to extend the safety net provision a little bit longer.” Several of our district informants explained that they did not think the safety net provision should be a permanent fixture; rather it would protect the students with disabilities who were in the middle or closing stages of their schooling and had not had sustained access to the general education curriculum:

The impact of raising standards in a system pre-K through 12 for students who are caught in between can be bad because they have missed what the early years could have prepared them better for. The children in special education who we’re dealing with right now in our middle schools and in our senior highs are kids who are in the between group and they need flexibility or it isn’t fair.

Informants in our large urban district also found problematic the state’s accountability requirement, under which schools only received credit for students who graduate within four years. This district had an initiative in place called the “Pathways Program,” which allowed students to take three, four, or five years to complete high school. As one informant pointed out:

We get to the commencement level and the kids take the test in 11th grade and then we draw the line at summer. In [our district] many of our cookies are still in

the oven. But it would be very simple for the state to say to the [districts like ours] of the world, your cohort analysis can go another year.

Another informant explained that when students with disabilities realized that their peers without disabilities were also taking an extra year to graduate, they felt better about themselves:

I think what that does is it helps kids who really do need the 5th year or the 6th year or the 7th year to feel better about themselves in taking longer. It changes the whole dynamics of what is expected and how long you're expected to be there.

However there was an inherent conflict between the graduation policy at the state level and the needs of students. The following statement is illustrative of many we heard concerning the mismatch between the state requirement for accountability purposes and the realities of this district:

The student is free and willing to go more years than they would ordinarily in a four-year program, but the school and the district are then accountable for every student that doesn't make it through in four years. So even though the student might be willing, you can see where teachers or administrators would feel that Pathways is a good thing, but there is a big disincentive to do it.

Changes in allowable testing accommodations

When we visited the state in October 2001, policy makers were in the process of revising the state's accommodations policy: "We just changed our policies to some extent on accommodations and modifications on the state tests." Our informants explained that the state recognized both accommodations and modifications and that the former did not invalidate the assessment score while the latter did: "An accommodation doesn't invalidate a test; a modification does, in our language." Our informants went on to

explain that the state assessments at the elementary and intermediate level were not high stakes for students and that schools needed to know where individual students were in terms of state standards:

The 4th and 8th grade tests are not high stakes, they are about determining where kids in a given school are, compared to standards. So if this kid can't read certain sections that will show up as him not being able to read certain sections, and his or her score will reflect this.

The same informant revealed that SED had thought long and hard about making these changes, but on balance believed that the change was necessary and would be beneficial to students with disabilities. As the following comment reveals, making this change would allow educators to see exactly where students were functioning, provide an incentive to districts to act quickly to remediate learning difficulties, and bring in academic interventions—which may be effective, especially at the elementary level:

Believe me, we had a long discussion about this, but in the end people felt, particularly in 4th grade, that if you allow accommodations to mask the kids' actual functioning then you will not know the nature of the problem and nor will the gate to academic intervention services be opened. Including those scores is essential as it takes away the incentive to allow a kid to get to that stage. You should be doing a lot more intervention early because you know the scores are going to count, and it also provides more accurate information about where the kid really is to the instructional staff and to the school.

We visited the districts in June 2002, just before changes in allowable accommodations were to take effect. As the following quotation illustrates, the district level informants were not particularly concerned by the proposed change. According to

this informant, developing a clearer policy on accommodations would make comparisons of student performance more reliable:

I'm not concerned about it. I don't think there has been consistency of how school districts have been using those accommodations across the state. So if across the state for these tests these accommodations aren't used then you can really do a better comparison of student performance. So having those taken away, those accommodations for those two tests, I'm not real concerned.

Addressing the assessment needs of students in the "grey area"

Informants at both the state and district level voiced concern for those students in special education who fall into the grey area. In the words of one state level informant: "People talk about kids in the grey area, between the two tests." These students were described as being too high functioning for the state's alternate assessment, but not high functioning enough for the state assessment. Another informant from the state explained: "There's these kids, and you don't know how many there are, that people anticipate are going to be scoring very, very badly, but yet they shouldn't take the alternate assessment." However, according to this informant, the state was still trying to decide what to do about the "grey group," as the state's alternate assessment was not appropriate for them:

The question will be down the road, what do we want to do? I don't see any of those kids moving into the alternate we have; the only ones we move in are the ones that should have been there in the first place. It's not as a result of doing poorly on something else that you are going to slide into that. They're very, very, very different assessment processes.

When we visited the districts in May 2002, the state had issued an interim policy to address the assessment of students with disabilities in the grey area. This interim policy, initially for one year, allowed districts to select another standardized assessment for “certain students with disabilities.” Informants at both our districts expressed concerns about the implications of this new policy. First, informants were concerned that the policy was a temporary fix for this problem and would only address the needs of a small number of students: “People are concerned about kids in the grey area, and talk about locally selecting assessments. Who would want to do this? It’s for one year and it’s for a very small population of kids that don’t qualify for alternate assessments.” Second, informants expressed concerns that high school students with disabilities may lose their eligibility for a local diploma: “Here’s the catch—if you’re a high school student, then you lose your eligibility for a local diploma. It’s very complex.” Third, informants did not believe that they should start to assess another group of students with disabilities separately from the general population: “I don’t think we should be testing even more students in a way that’s apart from the general population.” Finally, informants pointed out that the local assessment would not help them to determine where the student was performing against grade level standards:

We don’t plan to use it for very many students. Our feeling is, we’re going to do our very best to teach these students, and we need to see where they are performing, but if they don’t ever take the regular test, we would never find out. If we think it’s something that they can possibly manage, they do take it.

In this section, we have presented data from the state and district level relating to how well the state accountability policy was working for students in special education. Informants were positive about the importance of including this population in

accountability reform and described how many of the features of standards-based accountability reform—such as increased expectations, transparency of information, and use of performance data—clearly benefited students in special education and their parents. However, informant responses displayed some ambivalence about whether students with disabilities should be exposed to the full impact of accountability reform, especially at the high school level. The conflict among state policies, district policies, and the needs of students with disabilities also emerged clearly concerning high school graduation time limits. Informants also expressed concerns over the state’s interim policy for students in the grey area of the assessment system.

In the next section we present findings on the extent to which the state and the local districts have the instructional capacity to implement the Board of Regents’ ambitious standards-based reform initiative. In this section, we include informants’ comments about overall capacity to implement reforms as well as their capacity to implement reform for students with disabilities in particular.

Building Instructional Capacity

The success of standards-based reform depends heavily on the ability of all levels of the public education system to respond in appropriate ways to the information provided by accountability systems. In this section, we address the issue of instructional capacity across four related domains: (a) the state’s and the districts’ capacity to direct resources to support reform; (b) educators’ capacity to deliver the content, (c) educators’ capacity to teach students with diverse educational needs, and (d) campus principals’ capacity to be instructional leaders.

State and District Capacity to Support Reform

Financial Resource Limitation. The issue of how to best allocate valuable and scarce resources to support reform was discussed at both the state and district level. At the times of our state and district interviews both entities were experiencing significant budget shortfalls, which impeded their ability to fully implement the Regents' reform agenda. One participant at the state commented: "I think we are struggling with how to apply this, and with the resources that you have, how to make it work?" Another participant from the state described how her staff had been reduced from 8-10 individuals to just one:

Let me back up a minute, I have to tell you what our staff is. We have one math associate in our office. We used to have a bureau of math associates, 8-10 now there is one. So in our thinking this year of how we are going to leverage our resources and make sure that we can really have an impact, what can we do with one?

Another participant explained that the SEA was trying to realign their resources and working through the special education training and resource centers to provide content area information to special education teachers:

We are realigning our resources and now if you go to a Special Education Training and Resource Center (SETRC) you are going to get workshops on the math curriculum and how to modify it, new research-based instructional programs that will help teach it, and that kind of thing. So, it's just realigning all systems that support the school in order for the school to change.

Another participant explained that the SEA didn't have the resources needed to support the districts and that they had to rely on partnerships:

It's partnerships for us to try to get the message out, because we just can't do much with our lack of staff. I don't believe we can have a great impact on how districts roll these out right now, with the staff we have.

Another participant from the state explained that they had to choose between directing money toward improving math instruction or toward dropout prevention:

Here is some money that has just been proposed from the feds to our office, and we have two competing priorities that I met on this morning. One is dropout prevention and the other is the math initiative. I was explaining to the deputy commissioner, "let's invest in the math."

The issue of scarce resources also emerged at the district level, especially in our large urban district. When we conducted our interviews in May, this district had just heard that the superintendent needed to reduce staff across the district due to a severe budget shortfall. One informant discussed the implications of the layoffs on the district: "That's why this period of layoffs has been so painful for me. Because I see our investments going down the tubes and I know it's tough on everybody." According to another informant, the district was losing almost 1,000 employees: "Our budget deficit means that 946 positions are going, it's not going to be an easy piece, class sizes will swell and the discussion now is to push all the 6th grades back into our elementary."

This district, as one informant pointed out, served a very needy population and was able to qualify for extra federal resources:

In this district the poverty level is extremely high, it's about 80%, so we can get a number of initiatives up and running, but we're careful to make sure that we don't overwhelm schools, that they don't get too much and do nothing with any of it.

An unforeseen consequence of having a number of initiatives swirling around the school is the requirement that each department involved has to “prove” that the money is being well spent and that progress is being made. For example, we talked with two informants who worked in the district’s special education resource office. They explained that quite often they worked with schools that received multiple interventions funded from various sources. Each funding source needed data, usually the same data, to justify the monies spent. One participant commented that she spent more time writing reports than actually working in schools:

It’s really the same data and it’s going back to a number of different sources that are trying to validate the money that’s being spent. How many times do you justify the money that’s being provided to schools that you know need more resources? I understand to a certain point that you want to make sure that people aren’t wasting the money. The data collection process itself takes so much time away from the time that we’re actually out doing work in the school.

Human Resource Concerns. Like most states, New York was experiencing a shortage of certified teachers at the time of our study. As the following comment suggests, the shortage was expected to worsen rapidly:

I would say that around 13,000 teachers are currently uncertified. In 1998, we estimated that 50% of the workforce, which is around 220,000, were eligible to retire because they were 55 or over, within a decade. So in New York State we’re estimating that within the next seven years or so we’ll probably need 100,000 new teachers statewide.

Like many states, New York developed alternatives to the traditional certification routes in order to get people into classrooms. The target individuals are usually interested in a second career in teaching after a successful career elsewhere, and may be better able

than a recent college graduate to deal with working in a challenging environment:

Our intent is to take people that are trained, that have made it in another career, that are interested in changing. They are smart, they can pass the liberal arts and science test, and they know their content, but they're not kids, they are adults and they bring a different set of skills to the table than a 21-year-old entering a tough school. That's our hope of trying to widen the net in terms of capturing more people for the teaching profession.

In New York, the alternate route certification programs, like the traditional ones, are generally under the control of institutions of higher education. The reason for this is that the founders of the alternate route programs wanted to maintain the rigor of the certification program and believed the best way to do so was to keep higher education involved:

I was one of the architects of the alternative program, and we very precisely, very clearly, left it within the realm of higher education because we knew that we needed to keep them engaged and we knew that we had to demonstrate to a larger community that these are not second rate teachers—they're going to have the qualities you need, they're just going to get them in different ways.

However, as the same informant pointed out, there was an inherent tension in the Regents policy to increase both the quality and quantity of the teaching force: "It seems to me that the Regents were saying very forcefully, 'we're changing policy here,' and when they did this whole plan they understood that they had two main objectives and they were not necessarily compatible: quality and quantity." To deal with this tension, the Regents established the alternate certification programs, but also increased the accreditation requirements for all teacher preparation programs, required districts to provide one year of mentoring for all new teachers, and put in place procedures for

ongoing professional development. The following statement describes a few of the reforms made by the Regents:

The quality is the pre-service program, the one year of mentoring, the strong induction, the ongoing professional development, the more rigorous review. I mean all these things—we're requiring all our teacher education programs to be accredited by a teacher education accrediting body or the Regents. We require 100% pass rate on exams. Most of them make that.

Teacher shortage issues were experienced very differently by our two districts. In our more affluent and smaller district, informants reported that they did not have a problem finding new teachers and did not hire anyone who was uncertified or on a transitional certificate. As one informant commented: “ We are able to hire experienced teachers and pay for quality teachers because they want to work here.”

In our other district the situation was very different, as trained teachers left the district after a few years to go to the suburbs to get higher salaries: “We would train them and the suburbs would get them once we got them shaped up. They would be drawn away by salary, not because the city is too tough.” One informant told us that they had approximately 60 uncertified teachers in the district, mostly in mathematics, science, and special education. One informant addressed the issue of special education directly and explained that she expected a shortage due to retirements and the fact that the district, like many other big urban districts, was not very attractive to prospective teachers because of the financial problems:

The bottom line is that we don't look very attractive right now as a district, none of the big districts do, and we know we're going to lose some people over the summer, we know we have a certain number retiring. We have to advertise, but this year because of the problems in the district we've only interviewed a handful

of candidates in special education, so we know we will have a shortage.

This district had put in place two key initiatives that appeared to be successful in attracting new teachers and keeping them in the district, but these initiatives were in jeopardy because of the financial problems. The first initiative was an alternate certification program made possible by a partnership with a small college in the area and the second was a beginning teacher mentor program.

The district's alternate certification program, Urban Teachers for Tomorrow, was designed for professionals who were interested in a career in teaching, but were not able to go back to school to study for four years. Cohort 1, which was math, science, elementary, social studies and English, had over 40 individuals selected. Of the 40 who began, 34 people completed the program. The beginning teacher mentor program was also very successful, as the following quotation illustrates:

The new teacher retention rates I showed you, the average since the program began, is 86%, and in any medium-sized urban area with the poverty level that we have, anything over 65% is not bad. If I found out that this was down around 60%, this program would be too expensive to do.

Increasing educators' capacity to deliver content. It came as little surprise that the issue of educator capacity was very much on the minds of all state and many district informants. In many ways the data reveal little that is new on this issue, but rather underscore the pivotal importance of teacher quality to student learning. When we asked informants at the state level how schools could improve the performance of low-performing students and schools, the responses we received were similar to the following:

We've got to have the best teachers that we can possibly have because to move those kids around the first critical thing is the quality of the teacher. Our lowest performing schools in the state are typically ones where we have the least

experienced teachers and the highest turnover of teachers. If you don't have a good teacher there's no curriculum that's going to meet the needs of all those kids.

The Board of Regents had made changes to the teacher licensing examinations in light of the belief that they lacked rigor. One informant described the state's teacher education initiatives as part of the overall plan for increasing student achievement. The new examination requires initial certification teachers to pass the content specialty test:

In 1996, the Chancellor convened a task force on teaching to try to define what kind of teachers we want in our system to help ensure that all students moved to the higher learning standards. We will require in 2004 that pre-service teachers pass the liberal arts and science test, the assessment teaching skill written examination, and the content specialty test.

Another informant pointed out that the state hoped to improve the skills of both regular and special educators through additional changes in teacher preparation requirements, especially at the elementary level, and echoed the need for strong content preparation for special educators:

We have a sort of strategic overall goal, which obviously has to do with teacher capacity. It's really personnel capacity and we're working on it at different levels. For instance, you have the recent change in teacher preparation programs and that's aimed at preparing a much more content strong special education faculty. We are requiring all our teacher education programs to be accredited by a teacher education accrediting body such as NCATE, and 100% pass rates.

We asked our state and district informants what the main barriers to accessing the general education curriculum were for students receiving special education services.

Several informants identified the lack of content knowledge possessed by special education teachers as the most crucial issue. For example, this informant explained that the special education faculty members in many schools were not trained to teach content:

The barrier is special education folks not knowing what the curriculum is. But I don't think that's unusual. But you also have a lot of special education faculty and faculty in general who have been in schools a long time now and their training was not such that they were supposed to be able to teach math at a certain level or science at a certain level.

District level informants from our urban district concurred with the view presented above: "Practically speaking, and we've talked to a lot of special education teachers; it is math. It comes down to better preparation in mathematics for a lot of these special education teachers who are in the resource room." This view was shared by an informant responsible for content area training:

I find the people that I work with and do the training with are less comfortable with math. They're very, very comfortable with reading and writing, and listening and speaking that they're proficient at themselves and feel very competent at teaching someone else how to do.

Another informant from this district commented:

How many teachers know how to teach algebra to youngsters at the 5th and 6th grade level? They don't have that content background. The teachers are saying to me they don't have the content background or they don't have the instructional strategies to do this.

Our state level informants described several professional development programs that they had put in place. According to this informant regular educators, special educators and administrators all had to attend or the training wouldn't be offered: "The content training that we offer, always you have to have groups of regular educators with our special educators and our administrators or we won't offer it." However, according to one of our informants, there was plenty of money to fund professional development, but district superintendents did not always recognize the importance of continued professional development opportunities. According to this informant, the district board of education saw it chiefly as a way to placate the teachers' unions and stated that a new mindset was required:

If you added it all together it's probably hundreds of millions of dollars out there you could use for professional development. The key is showing the boards of education there's value in this activity. Until a superintendent does that then forget it. Many people consider professional development a sop to the unions so that teachers get the time off to do whatever they want to do. The mindset, I mean, when you talk about a 12 billion dollar industry, if you had a 12 billion dollar industry and continued development of your workforce is not part of it, you would not be a CEO for very long. But education works differently.

Because our two districts operated in very different socioeconomic contexts, they had different needs and adopted different strategies to increase content area knowledge for special education teachers. In our smaller, more affluent district one informant reported:

Last year we knew that this student would join us this year, so the special education teacher sat in on the regular education math classes and became familiar with the curriculum and expectations. She works with the students, and builds on the skills from regular math for some of them and teaches other students

who are in a more restrictive program. She is also going to get her endorsement in secondary math.

Increasing educators' capacity to teach students with diverse educational needs.

Informants at the state and district level pointed out that teachers may not have the pedagogical skills they need to teach children with a range of abilities. The following comment illustrates the view at the state level that all teachers need a variety of instructional strategies to teach the wide range of abilities found in general education classrooms, "It's critical that we provide teachers with real-world and intensive training so that they can serve any child in the regular classroom and not worry if they have a disability or not. That's number one." Another participant commented:

Many of the skills that we want teachers to possess are beyond academic areas. Part of what we want to do is make sure that all students are successful, so teachers have to be prepared to teach all students. Those teachers had to be prepared to teach children with special needs, they had to be prepared to teach children of low socioeconomic backgrounds, they had to be prepared to teach kids who were English language learners, and they had to be prepared to teach students of color.

In our urban district, informants were very vocal about the need for all educators to be trained to teach in urban settings such as their own. Our informants stressed repeatedly that urban environments brought with them many challenges, including extreme poverty: "An urban environment has a number of intrinsic challenges, not the least of which is students who are extremely poor, more poor than people often realize, and there are a lot of problems that are associated with that." Informants made it clear that teachers in urban environments needed a different set of skills than their suburban counterparts. The following quotation is illustrative of many we received on this issue:

The needs of high-poverty urban kids are different from kids in suburban middle class areas. We think that we should cram the learning standards that the state has said that kids need to be successful. However, we forget that they may be latchkey, they may be homeless, they may be all those things that kids deal with that we didn't have to deal with. Unless teachers are well trained within the affective side of what urban kids go through, we're going to miss it—it's almost like pouring water through a sieve.

Another informant from the district further explained that even with supports in place it was difficult to work in an urban school system. Teachers needed both to be committed to teaching in an urban district and to have a clear understanding of what that meant. If a teacher did not have both then he or she would find the job too hard:

There must be a commitment to teaching urban children and an understanding of what that means. If they do have a commitment to teach urban children and have some understanding of what that means, I think the mentor program and the administrators in this district can shape them up and make them fly. What we are finding is that they either have the commitment but not the understanding of what it is and we can't get them there, or they have the understanding but they don't have the commitment and so it's too hard, they want to go.

Increasing the capacity of campus and district leaders. Informants at the state level were very concerned about the need to increase the capacity of district and campus leaders to implement the Regents' reform agenda. One participant explained: "It's amazing how much leadership makes a difference, it really is, that's why you can find these schools all over the state and all over the country that shouldn't be doing very well, but are." As another participant commented, without good leadership a district could not respond even when the state went in to monitor it:

My sense is that there are districts that haven't gotten any better even with state monitoring because they lack leadership. No matter what else you do, you need at least competent leadership, and if you don't have that you are going to be struggling with those districts for a long time, because they can't respond without it.

The same participant explained that if a key leader left, then the district could quickly have problems:

One of the things we have found is that you could have a district that is doing well and the director of special education and the superintendent leave, and the next year they start to have problems because a new person comes in and doesn't have the same level of leadership skills.

The seriousness of the leadership issue at the school campus level was recognized at the state level. The following statement provides a picture of the stresses that principals face on a daily basis.

It is absolutely going to be as significant a problem long term because of the climate. It's not fun being a school principal. It's conflict with the boards of education, it's the remuneration level, it's the stress level, it's the leadership of the academic learning. It's all these different jobs.

Another state informant commented that the principals who she had spoken with felt overloaded and did not have the time to focus on student learning:

Look at our leaders, all of the pressure is on them right now. They have a million other things to worry about. I want to say wipe them off the table, here is where

we are, this is what you need to focus on. This is the curriculum: What are you doing and what are you teaching?

In this section, we discussed system level capacity to support the state's ambitious educational reform agenda. It is clear that policy makers at the state level and in our urban district especially, had concerns that they lacked both the monetary resources and human resources necessary to fully implement the Regents' plans. Policy makers at the state level were concerned about teacher shortages and also a looming shortage of campus and district leadership. Policy makers at both the district and state level recognized the inherent tension between quality and quantity and had established several initiatives designed to reconcile the situation. In our urban district, these initiatives were threatened by a budget shortfall.

In the following section, we address tension between the state's responsibility and the local districts' responsibilities regarding the public education system in New York. This tension emerged in the area of local curriculum alignment to the state's learning standards.

Alignment of District Level Curricula with State Content Standards

System cohesiveness is an integral component of standards-based reform and requires content standards, curriculum, and assessment to be closely aligned so that judgments on the performance of students and the effectiveness of schools are valid and reliable. If teachers are not delivering the required content, then students will not be successful on state assessments and schools will appear to be failing. The introduction of more rigorous content standards describing what all students in New York should know was a cornerstone of the Board of Regents reform initiative. However, when we visited

the state in October 2001, several informants stated that teachers were finding that implementing the state's new learning standards was difficult:

There are 28 standards, and each one of these has what we call performance indicators or the classroom level indicators for teachers to pay attention to. These weren't enough, people were saying, what does this mean, what does this look like, help me find it and see it in my class.

The state-level informants spoke emphatically on the importance of communicating the new learning standards to school districts and schools, but stressed that determination of what to teach in the classroom was firmly in the hands of the local school districts. Unlike some other states, New York did not adopt textbooks and had not developed a state curriculum, as the state believed that function should be the responsibility of the districts:

We are not a textbook state and we don't endorse textbooks and we don't put out lists of publishers who have aligned their curriculum to ours. Where we draw the line is developing actual state curriculum—that we are not going to do unless we are forced. We're still pushing back at the local districts saying no, this is your responsibility.

According to another state informant, the teachers' union in particular wanted the SED to provide much more guidance, even requesting a grade-by-grade breakdown. However, as the following quotation indicates, SED were not prepared to take on that responsibility:

The teachers' union kept asking us, please, give us the grade-by-grade, teachers need to know. But we refused. That is not our role, that is up to local school districts to determine what the curriculum will be. We will set the context of the

curriculum from which the district now can go in and align and match their local curriculum to this.

To help teachers become familiar with the new standards, the state department of education had developed a series of guidance documents for districts and schools. As the following two comments illustrate, many teachers and administrators thought that the new standards were just a passing phase. One informant stated:

In 1996, the teachers that came to hear about the standards didn't really believe we meant it, you know how it is, this is the fad, they come and go. But these documents went out to every area, to every school building, multiple copies went out.

Another informant said:

My fellow administrators and the union, their thinking was that this will not last. This is another requirement that's just not going to last. The commissioner said to them on day one when he walked into New York, we have three goals and three only, and in the past several years that has been his theme. Anybody today who thinks that we're going to be dropping this reform or backing up or slowing down, if they've listened to him then they know we're moving forward.

Although SED provided guidance documents in the content areas, our informants were uncertain of the extent to which teachers used them. One informant told us of a district where the guides had been in the district library for two years:

I was in one district and I see boxes of these sitting in the district library and I say to the librarian, do you think we could get these to the teachers? She said: "Of course we can." Okay, they have been sitting there for two years.

Our state level informants expressed considerable frustration that teachers were

not provided the opportunity to really understand the redesign of the content standards and to align their instructional content to the increased expectations. One informant noted: “They are not seeing the significance of taking those standards and bringing it down to the classroom level into instructional strategies for youngsters to learn and for teachers to know.” As this remark illustrates, until teachers understand how the learning standards relate to instruction, they will be resistant to change. As the following two quotations indicate, there was still a long way to go to get the field to support the changes. One person stated:

That’s where the rubber hits the road. Until we give teachers that opportunity to really sit down and see what it means and what it is going to look like, what are they to do other than say “This isn’t for me, it’s going to harm kids.” It’s just not real. So do I think it’s happening? I want to believe it is but I think we have so much more to do. If you ask an elementary teacher what the performance indicator is and what does that mean to you in terms of instruction, you are not going to see it too well. So that’s why I say we have a lot of work to do.

Another informant said:

There's learning it and doing it. I mean, we can hold these wonderful broadcasts on television, which is our strategy because we don't have the resources to get out to people, but unless they're doing, unless they're bringing student work to the table, unless they're bringing their own assessment plans and really getting into this and rolling up their sleeves, it's not going to happen.

One state informant commented that the issue of curriculum alignment to the state standards really came down to requiring teachers to take ownership of the reform process. If teachers were not involved in the nuts and bolts of implementation and remained

passive in the face of the state initiatives, their lack of involvement could become a way of avoiding culpability if students were not successful:

The union really wants us to come out and say on Monday we'll all be teaching this and so on and so on, and I keep thinking, how can you undermine these initiatives like that. So they are calling for a grade-by-grade curriculum so that if it doesn't work we know who to blame. They will push it away, push it away and then say, "It's not our fault."

When we conducted our interviews at the district level in May 2002, informants from both districts reported that their curricula were closely aligned with the state's revised learning standards. Indeed, according to informants from both districts, their respective superintendents had adopted standards-based reform before the state: "We had started down the standards path before New York State had because we kind of knew that that was where the universe was going to go. So then somewhere along in that process New York State standards came out." Comments from informants in both districts indicated that the process of curricular alignment was ongoing as the state continued to develop its content standards after the initial redesign: "The new standards gave us a clearer picture of our destination, so we had to realign after 1996 and do some benchmarking of our curriculum and standards with the new standards. We did it again in 1998-99 too."

Informants from both districts described several strategies they had used to bring about greater alignment. For example, there was a high rate of student mobility within our urban school district coupled with a range of campus-based reading curricula. The district decided to adopt a consistent elementary reading program for use in all 39 elementary schools. By adopting one approach district-wide, the Superintendent hoped to

provide a consistent approach to reading instruction regardless of whether the student stayed in the same school:

One of the very first things that the superintendent did address was this issue of the mobility of the students within our district and the impact that having the different reading series at the schools had on their ability to learn how to read. He immediately put into place an adoption for consistent reading programs throughout the 39 elementary schools.

However, as this informant pointed out, no single program could meet the needs of every student in every school. The approach in the district was to help teachers to use a variety of programs as resources to meet individual needs, instead of expecting one program to work with all students:

We paid attention to informing educators how to use the reading programs as resources in meeting the needs of children. I tell the teachers that the series isn't the reading program, you are the reading program. You're the professional, you're the informed person, you have the opportunity to meet the kids and to evaluate and assess on an ongoing basis what their strengths and their weaknesses are. Then you in turn pull from this toolkit and you employ what is necessary to get each and every child reading.

In the above section, we discussed issues around the alignment of local district level instruction with the state learning standards. At the state level, informants perceived that some districts and teachers did not accept that the Regents were serious about increasing educational standards and believed that if they waited long enough the initiative would disappear. There was also evidence that teachers did not fully understand how to implement the revised standards in their classrooms and did not develop a sense of ownership for the delivery of instruction. The role of the teachers' union was also

problematic, as it wanted the state to take over the role of districts in determining what should be taught, a role historically played by the local school districts, and one which the state did not wish to take on.

V. NEW YORK CASE STUDY DISCUSSION AND CONCLUSION

In the previous section, we presented qualitative data, collected over a two-year period—from two levels of the public education system, on the impact of accountability reform on students with disabilities in New York State. State level interviews took place in October 2001, and district interviews took place in May 2002. The sequential nature of data collection enabled us to study the impact of state accountability reform on students with disabilities and to follow the concerns identified by state level policy makers through to the districts. However, collecting data in this way meant that we were unable to obtain the initial views of state level personnel on the impact of NCLB, signed into law after the completion of the state level interviews, or to study the further development and refinement of federal accountability policy with this population. When we visited the districts in the following year there was some discussion of the impact of NCLB, but there was not a clear understanding of how the legislation would affect students with disabilities. EPRRI staff interviewed 29 individuals, 8 at the state level and 21 at the district level. Three broad themes emerged from the data: (a) making accountability work for students with disabilities, (b) building instructional capacity, and (c) curricular alignment.

When considering the impact of standards-based accountability reform on students with disabilities in New York State, it is important to remember the newness of these reforms. Before 1996, the state's involvement in education was concentrated at the

commencement level examinations, with local districts having complete discretion over what was taught, how it was taught, and to whom it was taught. This situation changed rapidly over a short time period, with greater involvement and increasing requirements from the state, which was followed by federal requirements under NCLB. When we visited the state in October 2001, we found a system in transition with the concomitant fears and speculations that frequently accompany such changes. With this mind, the following key findings emerged from this study.

First, New York State was unique in respect to the Regents' commitment to include students with disabilities in educational accountability reform from the beginning. Informants at the state and district level commented that reporting the performance of students with disabilities was particularly important, as it challenged the low expectations of parents and teachers, empowered parents to challenge their local education leaders, brought the poor performance of this subgroup to the attention of district superintendents, and provided examples of schools and districts that were beginning to improve the performance of this population. Comments from our state level informants indicated that before the Regents required public reporting and disaggregation of the scores of students in special education, no one had any idea how badly this population was performing. Comments also indicated that some district superintendents and boards of education were very not very pleased, possibly due to the negative portrayal of their districts, but eventually started to pay more attention to the instructional needs of this population.

However, despite the intention of the Regents to include students with disabilities in educational reform, the practical implementation of this policy was problematic,

particularly once NCLB required comparable improvement. This population of students in special education did not fit neatly into the theory of action behind accountability reform, especially as the Board of Regents had adopted very challenging performance standards, assessments, and graduation requirements. This lack of fit necessitated that the Regents allow additional flexibility for students with disabilities in the form of the safety net at the commencement level and the local assessment provision for students with disabilities in the grey area. This flexibility allowed the system to treat students with disabilities differently from other students, particularly when the stakes were high for students and schools.

A second finding is that policy makers at the state and district level reported that the performance of students with disabilities was slowly improving, with higher graduation rates and evidence that students were moving at least from a Level 1 to a Level 2 in the performance index. However, informants in our urban district in particular acknowledged that there was a long way to go for this population as well as others. Interestingly, several informants reported that students with disabilities in some schools were performing better than their peers without disabilities. However, this was not necessarily a cause for celebration, as it seemed to merely reflect the very poor performance of general education students overall.

A third finding is that any reform effort requires system capacity to support effective implementation of policy. The lack of financial capacity at the state and in our urban district placed the success of the Regents' educational reforms in serious jeopardy, especially the goal of reducing the achievement gap. Over half of the state's students are enrolled in high-need districts, including one of our study districts, and as student poverty

in a school increases, academic performance tends to decline. Districts with needier students tend to spend less because they have limited fiscal capacity due to local tax effort difficulties.

State level informants did not have the staff or resources necessary to support the local school districts in aligning their curricula with the state's learning standards, or to help teachers change their instruction in ways that would enable students to reach the higher standards. State level informants perceived the lack of alignment as a serious obstacle to education reform; however our two participating districts reported that they had adopted standards-based reform before the state and had simply tweaked their existing curricula. According to our state level informants, not all districts were in the same position.

In addition, state informants also reported that district boards of education did not recognize the need to provide high-quality professional development and that this was a significant barrier to raising the quality of existing teachers and leaders. State informants emphasized the crucial importance of leadership at the district and campus level in turning around struggling districts and schools. In our urban district, informants made several comments about the need for commitment on the part of the district superintendent to increasing the performance of all students. Several informants expressed their concerns and fears of a looming shortage of school principals due to the increased pressure that principals faced.

New York, like most states, was trying to solve two capacity needs that were not necessarily compatible: increasing the number of teachers and increasing the quality of the teaching force. In the state of New York, the Board of Regents exercised control of

education policy in the public school system and also in higher education, which enabled them to address these issues on two fronts. Thus, the Regents increased pre-service teacher education standards in content and instruction, required all teacher preparation programs to be accredited and to have 100% pass rates on the state's licensure examinations, and increased control over professional development requirements.

In terms of content area skills, state informants reported that special education teachers needed to increase their ability to teach mathematics and that general education teachers needed to improve their pedagogical skills to be able to teach all students. Informants from our more affluent district reported few problems in terms of teacher quality or quantity, but the situation was very different in our urban district, where informants reported significant numbers of uncertified teachers. Informants from this district emphasized that teachers who worked in urban environments needed a special set of skills, including an understanding of the impact of poverty on students and a commitment to teaching in the urban environment.

A fourth finding concerns the issue of stakeholder commitment to education reform. State informants perceived the support from the New York State Teachers Union and some districts and teachers to be lukewarm at best. Several informants at the state level perceived the union to be unconvinced of the likely staying power of the Regents' reform agenda. The union wanted the state to provide a grade-by-grade curriculum for the districts, an action which the state did not want to take. State level informants wanted to maintain the philosophy of local control of the curriculum within the framework provided by the learning standards, and also to engage teachers in the reform efforts. One informant feared that the lack of engagement on the part of teachers and the union might

be a way of abdicating responsibility for reaching the higher standards.

In conclusion, when we conducted state-level interviews in October 2001 and district level interviews in May 2002 standards-based education reform was still in the early stages of implementation and the extent of teacher buy-in to the reform agenda varied from district to district. An interesting feature of the education reform initiative established by the New York Board of Regents was that they explicitly included students in special education, but also allowed flexibility for this population in the form of the safety net when the impact of high stakes testing fell directly on the student, as it did at the commencement level.

Initially, the Regents' fledgling accountability system was applied only to Title I schools, allowing non-Title I schools to operate under the accountability radar. NCLB changed the situation by introducing high stakes for all schools and requiring comparable improvement by specific target groups as measured by the state's general assessment. New York State had adopted a set of challenging performance standards for all but the most severely involved students with disabilities, and it was against these performance standards that most students with disabilities were being measured and all schools were being judged. This sudden ratcheting up of educational standards created significant tension between the inclusiveness of the Regents' policy and the requirement that all schools make AYP against challenging academic performance standards.

One result was the emergence of a "new" group of students in special education whose academic progress could not be measured by either the general assessment or the state's alternate assessment. New York State initially developed an interim policy allowing districts to select an assessment, subject to certain criteria, and planned to

include these students in the school accountability system as Level 1 participants. When we talked with informants at our two participating districts in May 2002 they were not enthusiastic about the local assessment option, but this view may have changed when the full impact of NCLB was better understood and districts realized that this population could impact AYP status significantly. Once the U.S. Department of Education began to respond to concerns from the field about the performance of students with disabilities and their impact on AYP, SED indicated it would make full use of any additional flexibility granted by the federal government.

Our findings from New York State clearly indicate that the theory underlying accountability reform can work for students with disabilities, but only up to a point. The theory assumes that publicizing a school system's achievement outcomes will make the system work harder and that students will work harder when promotion and graduation are at risk. Informants at both the state and the district level provided some evidence of the success of accountability reform for this population, but also indicated many factors that interfere with the workings of this model, factors that the model ignores. These factors include (a) a lack of money to support system alignment, (b) a lack of teacher capacity to improve instruction, (c) questionable grassroots support of the reform agenda, and (d) the difficulty of introducing flexibility into a theory that by its very definition is inflexible—holding all students to the same challenging requirements.

In this case study, we have presented our findings on the impact of accountability reform on students with disabilities. Data were collected at two different time periods from two levels of the New York State public education system. The policy context changed dramatically during the period of study. The locus of control shifted to the

federal government at a time when the state had only just put the finishing touches on its own accountability policy, in which students with disabilities figured prominently. When we began, the state's direct involvement in education was recent and fairly limited in that SED focused on outcomes—what students should know and be able to do. SED wanted to keep its distance in terms of curriculum and instruction, leaving the “how” to the districts. In addition, accountability was focused on a subset of schools in the state—those identified as Title I schools—and covered a restricted number of grades. The implementation of NCLB required the state to make substantial changes to its fledging system before the local school districts had really had a chance to understand it. Moreover, NCLB was far more expansive than New York State's accountability system in that it covered all schools and all students, including targeted student groups, which consequently caused a considerable increase in the number of state assessments.

References

- Bogdan, R., & Biklen, S. (1992). *Qualitative research for education: An introduction to theory and methods*. Needham Heights, MA: Allyn & Bacon.
- Cohen, M. (2002, February). *Implementing Title I standards, assessments, and accountability: Lessons from the past, challenges for the future*. Paper presented at the conference sponsored by the Thomas B. Fordham Foundation, Washington, DC.
- DeBray, E., Parson, G., & Woodworth, K. (2001), Patterns of response in four high schools under state accountability policies in Vermont and New York. In S. H. Fuhrman (Ed.) *From capitol to the classroom: Standards-based reform in the States* (pp. 170-192). Chicago: University of Chicago Press,
- Eisenhardt, K. M. (2002). Building theories from case study research. In A. M. Huberman & M. B. Miles (Eds.), *The qualitative researcher's companion* (pp. 5-36). Thousand Oaks, CA: Sage.
- Eisenhardt, K. M., & Bourgeois, L. J. (1988). Politics of strategic decision-making in high velocity environments: Towards a mid-range theory. *Academy of Management Journal*, 31, 737-770.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies of qualitative research*. London: Wiedenfeld and Nicholson.
- Individuals with Disabilities Education Act Amendments of 1997*, 105-17, 20 U.S.C. § 1401(a)(17).
- Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousands Oaks, CA: Sage.

- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.
- Pettigrew, A. (1988). *Longitudinal field research on change: Theory and practice*. Paper presented at the National Science Foundation Conference on Longitudinal Research Methods in Organizations.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.
- Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago: University of Chicago Press.
- Yin, R. K. (1989). *Case study research: Design and methods*. Newbury Park, CA: Sage.