Raising Standards in American Schools: the Case of No Child Left Behind

Smith, Emma

DOI:
10.1080/02680930500132403

Raising Standards in American schools: the case of *No Child Left Behind*

Emma Smith
Department of Educational Studies
University of York
YO10 5DD
Email: es25@york.ac.uk

Abstract

In January 2002, President George W Bush signed into law what is arguably the most important piece of US educational legislation for the past 35 years. For the first time, Public Law 107-110 links high stakes testing with strict accountability measures designed to ensure that, at least in schools that receive government funding, no child is left behind. The appropriately named *No Child Left Behind Act* (NCLB) links government funding to strict improvement policies for America's public schools. Much of what is undertaken in NCLB is praiseworthy, the Act is essentially equitable for it ensures that schools pay due regard to the progress of those sections of the school population who have traditionally done less well in school, in particular, students from economically disadvantaged homes, as well as those from ethnic minority backgrounds and those who have limited proficiency to speak English. However, this seemingly salutatory aspect of the Act is also the one that has raised the most objections. This paper describes the key features of this important piece of legislation before outlining why it is that a seemingly equitable Act has produced so much consternation in US education circles. Through an exploration of school level data for the state of New Jersey, the paper considers the extent to which these concerns have been justified during the early days of *No Child Left Behind*.

What is No Child Left Behind?

> 'There have been two basic policy eras in US education policy since mid-century: a struggle for access and equity that dominated the period from 1960 to 1980 and a

'This is an electronic version of an article published in The Journal of Education Policy *Vol. 20, No. 4 (July 2005): 507-524.* The Journal of Education Policy *is available online at: http://www.tandf.co.uk/journals/titles/02680939.asp.*

focus on competition and standards that prevailed in the 1980s and 1990s' (Orfield 2000, p406).

This re-focusing of US education policy was perhaps no better demonstrated than when the *No Child Left Behind* legislation was passed into law by President George W Bush early in 2002. For a seemingly equitable Act, which endeavours to ensure the academic progress of *all* students, the levels of criticism that have met its inception have been surprising. Despite this apparently equitable intent, some commentators fear that the high stakes testing and accountability-linked sanctions that underpin the Act could result in many otherwise successful schools being labelled as failing. Through an examination of how strict accountability measures became so entwined with government policy to raise standards in American public schools and an exploration of school level data for the state of New Jersey, this paper considers the extent to which these concerns have been justified during the early days of *No Child Left Behind*.

No Child Left Behind requires that all schools and school districts which receive Title-1 federal funding put into place a set of standards for improving student achievement, together with detailed plans charting how these standards with be monitored and met. Title-1 is a part of the Elementary and Secondary Schools Act which distributes federal funds to disadvantaged areas, about 90% of America's 15,000 school districts receive Title-1 funding (Ravitch 1995). A major consequence of these standards is that schools will be required to set targets and monitor the progress of students, and subgroups of students, in order to ensure that 100% of *all* students reach certain minimum proficiency levels by 2014. Failure to achieve proficiency would lead to 'corrective action', which, in its most extreme manifestation would result in school closure (Department of Education 2002). Unlike the UK, the US has a very de-centralised system of education, with much of the control over schools devolved to school districts which act on behalf of the state. There is no national assessment system nor a national curriculum, the responsibility for ascertaining standards, assessment tools and curriculum coverage lies with the individual

states. However, just like in the UK, a 'crisis account' exists over the apparent underachievement of American schools, particularly with regard to their relative performance in international comparative tests (Schmidt *et al* 1999). Legislation like *No Child Left Behind* is designed, through complex systems of school accountability and sanctions, to remedy this.

The concept of accountability coupled with high stakes testing is not new in US educational policy; indeed NCLB is a composite of earlier legislation, including state-wide accountability protocols and testing regimes. During the 1990s, the majority of states introduced an element of state-wide testing and, to a lesser extent, accountability measures, and by 2002 most had some form of testing procedure in place (Rudalevige 2003). However, what is new, is the scope and potential impact of the NCLB sanctions, and many commentators fear that the undertaking to ensure that every child reaches full proficiency may result in large numbers of otherwise successful schools being labelled as failing (Kane and Staiger 2003, Popham 2004).

By January 2003, each state was required to submit to the US Department of Education a detailed workbook outlining the steps they would undertake to ensure compliance with the statutes set out under NCLB. The Act itself requires that by 2005, states assess performance annually in grades 3 to 8[1] in Language Arts Literacy and Mathematics, and in Science by 2007. Additional tests must also be administered to students during grades 10 to 12[2]. States must also indicate how both schools and school districts will demonstrate Adequate Yearly Progress (AYP) towards full proficiency by 2014 and make public their test results. This performance data will also be disaggregated according to different student sub-groups, characterised by students' sex, minority group, Special Educational Need, level of economic disadvantage and English language proficiency. Typically, a subgroup would comprise 25 students, although states do vary in their

[1] National Curriculum years 4 to 9
[2] National Curriculum years 11 to 13

definition of this (Department of Education 2002, New Jersey Department of Education 2004). In addition to annual testing, schools must also ensure that at least 95% of all students are assessed. If this target is not met, then schools will not make AYP, regardless of the proficiency scores of the remainder of the cohort (Popham 2004). If a school fails to make AYP, a series of sanctions can be administered by the school district. The form of these sanctions ranges from district level monitoring through to giving parents the option to transfer their children out of 'failing' schools and providing students who remain in the school with additional tutoring. In more extreme cases, where a school fails to make AYP for four or more consecutive years, that school can be faced with having to replace staff, aspects of the curriculum or, at the extreme, be re-structured as a Charter school or one run by a private company.

'Compassionate conservatism' and the 'failure' of public schools: The origins of NCLB

The NCLB Act has risen from a 'primeval soup' (Kingdon J., in Rudalevige 2003, p27) of education policy to raise standards in America's public schools that has spanned several decades. Although many of the components of the Act contain little that is completely new, what is unusual is how the Act managed to achieve widespread bipartisan support from Congress. Traditionally the federal government, and Republican administrations, in particular, has avoided much direct influence in educational initiatives, particularly as federal funds only contribute about 7% of a state's total educational revenue (Hochschild 2003). In order to appreciate this unprecedented level of federal interest in educational achievement in the context of the *No Child Left Behind* Act, it is important to recognise that standards in America's public schools have long been under scrutiny. The launch of Sputnik in 1957 by the USSR had far reaching repercussions and created pressure on schools to raise academic standards, as well as enrolment on mathematics, science and foreign language courses (Ravitch 1995).

Subsequent decades of falling or stagnating scores on two key nationally administered tests, namely the Scholastic Aptitude Test (SAT) and the National Assessment of Educational Progress (NAEP), coupled with dubious performance on international comparative tests, reaching as far back as the First International Maths and Science Study (FIMSS) in the 1960s, contributed to the publication in 1983 of a searing indictment of educational standards when the Regan administration released *A Nation at Risk*. The invective used in this document is strong and condemns the 'rising tide of mediocrity' (NCEE 1983) which was eroding the American public school system:

> 'If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war' (NCEE 1983, p3)

While the emphasis on raising standards can be traced back to at least the 1980s, federal interest in public school accountability, coupled with high stakes testing and elements of school choice has been evident in education policy reforms ever since. Consequently, the *No Child Left Behind* Act 'collected and encompassed proposals advanced in theory and substance for years, accrediting Ronald Regan-, George H W Bush- and Bill Clinton-era initiatives into a single bill' (Rudalevige 2003, p24).

Having established the key developments surrounding the inception and content of this new piece of legislation, the following section will consider how some of the accountability measures demanded under *No Child Left Behind* are working in practice during the early days of the Act.

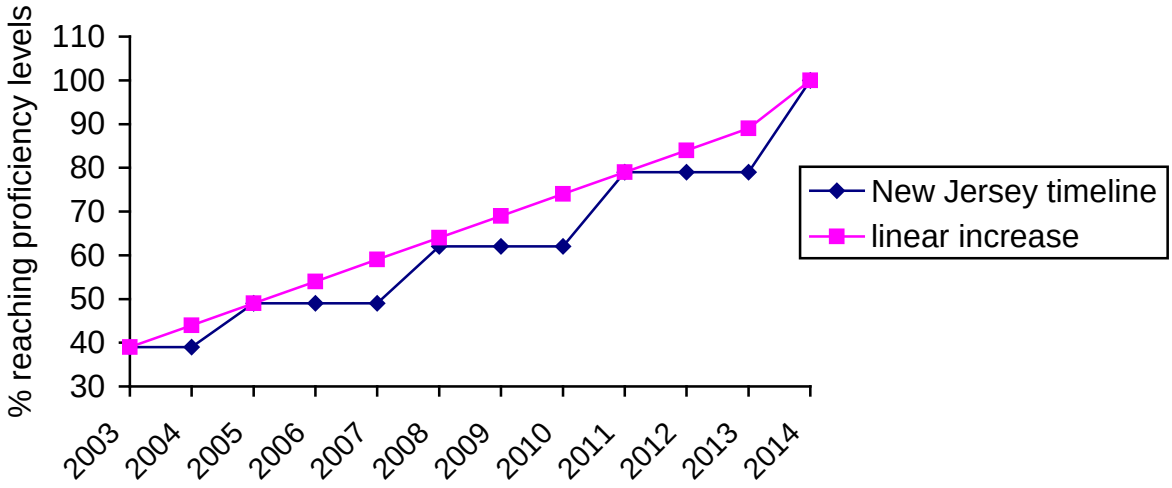
Making Adequate Yearly Progress: concerns over rules for student subgroups

On the one hand raising the achievement of students in all America's public schools would appear to be both equitable and praiseworthy. However, some of the implications

of NCLB's strict accountability rules and sanctions, in particular those that apply to student subgroups, could result in great numbers of schools and their students being labelled as failing. One of the strengths of the Act is that, as its very title suggests, it demands that the academic progress of every child, regardless how able, be open to scrutiny. However, the reason that many commentators and practitioners take issue with this, is that the Act also states very clearly that not only is every child expected to make progress, they must make sufficient progress to achieve minimum competency levels within 12 years of the Act's inception. According to Linn, one of the flaws with the 100% proficiency target is in its expectation that all schools must achieve the 100% levels, even though the number of schools who are actually at those levels today is very small. In other words, he argues that 'we should not set a goal for all schools that is so high that no school has yet achieved it' (2003, p4).

As NCLB does not mandate specific annual progress targets towards full proficiency, every state is required to chart their own timeline for making Adequate Yearly Progress, so that by 2014 each student, and student subgroup, achieves at or above the state's proficiency levels. The Act, however, does specify that students must make annual *incremental* progress towards full proficiency. The temptation, of course, is that states will set their AYP targets very low, focusing for example on basic skills tests (Hess 2003, McNeill 2000, Haney 2000). According to Popham (2004), such incremental progress would mean that many schools would have to raise the number of 'at proficiency' students by 5 or 6 percentage points annually. For these schools, sustaining such year on year increases could result in many of them failing to meet AYP targets within a few years. Indeed, according to Lee, the progress rates of many schools would have to increase by 6 or 7 times if proficiency targets are to be reached by the 2014 deadline (Lee 2004). In their efforts to minimize the chances of costly 'AYP-induced failure', some states have crafted what Popham calls 'inventive' timelines. Figure 1 below shows the timeline for achieving full proficiency in state tests in Grade 8 mathematics in New Jersey.

Figure 1: New Jersey's incremental timeline for AYP in Grade 8 mathematics



Source: New Jersey Department of Education 2004

'Inventive' timelines such as these are designed to reduce the likelihood of schools failing to make AYP during the early years of the Act. For example, in New Jersey in 2005, 49% of students are expected to meet proficiency levels in Grade 8 mathematics. Schools, in order to make AYP, would first have to achieve and then maintain these levels for the next two years before making a relatively large 13 percentage point jump in 2008, when 62% of students would have to be proficient. This means that between 2003 and 2007, school proficiency levels, on this scale will only have to increase by 10 percentage points, rather than the 20 percentage points required on a linear scale.

It also means that only about half of the school's students have to be proficient by the end of 2007, leaving the remaining 50%, presumably students of lower ability, or those with Limited English Proficiency, to achieve these levels in the remaining 7 years of the Act. Additionally, in 7 of the timeline's 12 years, no annual progress at all is required. The justification for leaving larger jumps in the numbers of students achieving proficiency levels until later on in the Act's lifetime would seem sensible. It is a new and untried

piece of legislation and the penalties for not achieving AYP are potentially stringent, it is therefore very likely that reform to the Act will take place, perhaps in the form of softening accountability sanctions (Hess 2003). One need only read an edition of *Education Week* (the US equivalent of the *Times Education Supplement*) to see that amendments to NCLB legislation are likely. It should also be noted that New Jersey is by no means alone in setting timelines like this; Ohio, for example, has a similar system (Linn 2003).

However, even at the start of the 12-year timeline towards full proficiency, the numbers of schools failing to make AYP are large. In Washington State, in 2003, 436 out of 2000 schools failed to make AYP (Bylsma 2004), predictions for California for the 2004-05 school year indicate that almost two-thirds of the state's schools will not make AYP (Perry 2004), and in Missouri half of the state's 2000 public schools failed to make AYP in 2003 (Education Week 2004b).

Another important feature of these AYP timelines is how they compare with the states' own accountability measures. Several states already have their own accountability targets, many of which have been in place for some time. In fact 21 states are maintaining their own accountability systems in parallel to NCLB (Education Week 2004a). However, the two systems do not necessarily concur, in fact, in just about every state a higher proportion of schools met state AYP targets but not targets for NCLB. For example, in California, 78% of schools met the state accountability requirements, compared to only 55% for NCLB, and in North Carolina, 90% of schools met the state requirements and only 47% made AYP under *No Child Left Behind* (Education Week 2004a). This presents a dilemma, not only for parents whose children's schools, on one accountability measure are successful and on the other, are labelled as 'failing' and subject to corrective sanctions. It certainly begs the question of whose accountability system most accurately reflects progress and achievement in schools. It would seen bizarre, but certainly not unlikely to foresee a scenario where a school that has failed to make AYP under *No Child*

Left Behind for two consecutive years is having to offer students the choice of attending another local school, while on the state's accountability measures, that school is a success.

One of reasons why many schools which would otherwise be successful on state accountability measures, might be deemed to be failing under NCLB, is due to the Act's strict rules for student subgroups. While few would disagree that challenging the progress of groups of students who have traditionally done less well in school is praiseworthy; the issue lies with the expectation that these students must also meet the demanding AYP proficiency targets otherwise their schools will receive stiff penalties. According to Kane and Staiger, the use of subgroup rules is 'counter productive in test-based accountability systems' (2003, p152), while Linn argues that 'the goals that NCLB sets for student achievement would be wonderful if they could be reached, but, unfortunately, they are quite unrealistic, so much so, that they are apt to do more to demoralise educators than to inspire them' (Linn 2003, p10).

That 100% of students are required to reach full proficiency levels, when relatively few do at present, is to some commentators unworkable, not least because of the way that disadvantaged groups of students are unevenly distributed in America's schools. These are concerns that focus mainly, although not exclusively, on the educational experiences of children who live in America's large cities. Although the United States is a wealthy country, there exist within it large pockets of poor and isolated groups. With one of the most unequal distributions of wealth of any industrialised country, the US has large sections of its population living in poverty, mainly in the urban areas of large cities (Orfield 2000). The fact that students from different economic and cultural backgrounds are not distributed evenly throughout the USA, and indeed are clustered in certain localities, has important implications for the NCLB subgroup rules. With many commentators suggesting that the sanctions linked to the subgroup proficiency measures will result in diverse schools and schools with large numbers of students being unfairly

penalized (Abedi 2003, Kane and Staiger 2003, Popham 2004, Lee 2004). In the USA, these schools are likely to be those with large proportions of students with Limited English Proficiency (LEP), as well as children from certain ethnic minority backgrounds and those from poorer homes.

At present in the United States, over 7 million children are enrolled in elementary and secondary schools in the nation's central cities, and almost a third of all children of school age (over 15 million children) live in the suburban fringes of these cities (Department of Education 2002a, NCES 2002). However, the distribution of children within both these communities is uneven. Only 6% of white children attend schools in urban areas, compared with 31% of children from ethnic minority backgrounds (in particular African-American and Hispanic children) (Department of Education 2002a). Nationally, 16% of children, between the ages of 5 and 17, are identified as living in poverty, 24% of these children live in the central cities, while 10% live in their suburbs (Department of Education 2002a). In some states like New Mexico and California, over 80% of schools contain a Hispanic or African-American subgroup, compared to only 5% of schools in Virginia. Additionally, 92% of African-American students and 91% of Hispanics attend schools with Black or Hispanic subgroups, compared with only 33% of white students (Kane and Staiger 2003).

That African-American students perform less well on standardised tests than European Americans is well documented in US educational research. Longitudinal assessments since the 1960s have shown white students to be ahead in every measure and at every grade (Jencks and Phillips 1998). More recently, data from the National Assessment of Educational Performance (NAEP) revealed wide disparities between the attainment of African-American and white students. Similar trends were also reported between the performance of students from white and Hispanic backgrounds. According to the authors of the study, 'the single most important determinant of the difference in failure rates between states is likely to be the racial composition of their schools' (Kane and Staiger

2003, p175). In addition, the reason why almost 57% of Ohio's school districts failed to make AYP in 2003 was either partly or fully because of the progress of students with SEN (Chester 2004). In Washington state, it was larger schools with more diverse populations who were also less likely to achieve yearly accountability targets (Bylsma 2004). Texas, one state held up as a model of school accountability reforms, has for some time monitored proficiency targets for subgroups of students. However, the expected targets were different for students at aggregate and disaggregate level: 90% proficiency targets overall and 55% proficiency targets for subgroups of students (Kane and Staiger 2003).

So on the one hand, we have a seemingly equitable piece of legislation, designed to give all students equal chances of success in school. However, schooling in America's public schools appears to be anything but equitable. The concern is that, in its present form, the *No Child Left Behind* legislation could further enhance these inequities and unfairly label children and schools as underachieving. In the final section of this paper, we consider the extent to which some of the fears of the opponents of *No Child Left Behind* have been realised during the early days of the legislation. Using the state of New Jersey as a case study, we consider how accurately we are able to label schools that do not make AYP as underachieving or failing.

No Child Left Behind - New Jersey style

New Jersey is the most densely populated state in the USA. It has a population of around 8 million people, only 10% of whom live in the New Jersey's six largest cities. In 1999-2001, the median household income in New Jersey was above that for the rest of the nation ($52,137, compared with $42,873 nationally), and fewer children and adults were identified as living in poverty (NCES 2002). New Jersey has 626 school districts, 120 of which are classed as wealthy and 30 districts that are designated as the poorest or 'special needs' districts. These poorest districts are also known as the 'Abbott' districts, so named after the plaintiffs who brought about the legal challenges to the distribution of school

financing in New Jersey during the 1980s. The New Jersey Department of Education categorises school districts in the range A-I based on their relative wealth and socioeconomic status. Schools in the A category are overwhelmingly, although not exclusively, located in the poorer Abbott districts, while schools in categories I and J are located in the wealthiest school districts (New Jersey Department of Education 2004b). This categorization of New Jersey's school districts is important, educational policy in New Jersey has been shaped by over 25 years of school finance litigation and by 20 years of state efforts to hold school districts financially accountable for the quality of the education they provide. In June 1990, the state supreme court declared that New Jersey's school finance law was unconstitutional, that urban schools were not providing their students with the 'through and efficient' education demanded in New Jersey's constitution and that funding inequities should be eliminated (Firestone *et al* 1997). In its ruling, the New Jersey state supreme court required that more financial resources go to poorer urban districts than to wealthier ones. The premise for this was that because of their additional social and community responsibilities, schools in the urban districts needed more money than those in wealthier areas. Schools in urban New Jersey had suffered many decades of under-funding and decline, not just because of changing demographics in the central cities, but also because of the way in which education is financed in the USA (Anyon 1997). In America, public school funding is available from three main sources, the federal government, who typically contribute about 7-8%; the state (contributing around 48%) and local taxes (around 45%), although the exact proportions do vary by state (Adams and Adams 2003, Flanagan and Gissmer 2004, The Education Trust 2002 and table 1 below). States tend to delegate their proportion of the funds based on student numbers and student characteristics and under this mechanism urban areas do tend to get more funding. However, local financing tends to be closely linked to property taxes. Therefore higher property values will raise more money for local schools. In urban areas, where property prices tend to be lower than in the suburbs, the funds raised for schools by property taxes are proportionally lower than those raised in suburban districts. These different funding mechanisms result in many of America's

urban districts being funded inequitably. For example, in New York state, where the funding gaps are among the largest, pupils who attend school in districts in the upper quartile for poverty can be allocated $2,152 per student less than if they had attended a school in a district in the lower quartile for poverty. These $2,152 equate to a difference of $53,800 for a class of 25 students, or $860,800 for a school of 400 students. Similar gaps in state and local funding appear between districts with the highest and lowest proportions of students from ethnic minority backgrounds; a gap which nationally corresponds to $902 per pupil, or $22,550 for a class of 25 students and $360,800 for a school of 400 (The Education Trust 2002).

In New Jersey, where school funding is among the most equitable in the whole of the United States, mainly as a result of judicial intervention over the last two decades, students in the poorest school districts do receive a greater proportion of state funding and per pupil expenditure than those in the wealthiest districts (table 1). Notice also the relatively low proportion of funding that is provided by the federal government, typically around 3%. It is this funding stream that is directly linked to the NCLB sanctions.

Table 1: Distribution of school funding according to the relative wealth of school districts in New Jersey.

District	Number of schools	Mean revenue from local sources (%)	Mean revenue from state sources (%)	Mean revenue from federal sources (%)	Per pupil expenditure ($)
A	398	13	75	5	13227
B	263	41	49	4	10985
CD	229	52	39	3	10470
DE	349	62	30	3	10089
FG	293	69	24	2	10588
GH	306	78	17	2	11083
I	405	85	11	1	10962
J	35	86	6	1	11646
Total	2278	57	35	3	11156

Source: New Jersey Department of Education 2004.

New Jersey is an interesting example of a state that has sought to eliminate inequities in school funding, while attempting to ensure adequate for provision for students who attend schools in the poorest districts. How the NCLB accountability targets impact upon the State's attempts to provide the 'through and efficient' education demanded under its constitution will be considered below.

State testing and school accountability

Since 1978, students in New Jersey have been assessed in grades 3, 6 and 9 or 11 in reading and mathematics. Over the past two decades, the testing regime has been revised and updated to include basic skills tests, curriculum content standards, accountability measures, public reporting of district level scores and high school graduation tests (Firestone *et al* 1997). When NCLB was implemented in 2002, students in New Jersey were being assessed state-wide in Grade 4 (The New Jersey Assessment of Skills and Knowledge 4), in Grade 8 (the Grade Eight Proficiency Assessment) and in Grade 12 (The High School Proficiency Assessment). In 2003, 208 schools, or around 10% of New Jersey's elementary and high schools, were identified as being "in need of improvement" for a second year (New Jersey Department of Education 2003). These schools did not meet all their Adequate Yearly Progress targets despite the fact that many did reach or exceed their targets in one of two content areas. The consequences of failing to make Adequate Yearly Progress for two consecutive years mean that parents are now free to transfer their children to other schools in the district, while the school will continue to receive technical assistance with its curriculum and teaching programmes.

In the following sections we look more closely at the key characteristics of these schools. By comparing these 'failing' schools with other schools in New Jersey, and by taking into account contextual features of these schools, we consider the extent to which the sanctions demanded under *No Child Left Behind* can be justified.

The characteristics of New Jersey's schools

The data used in this analysis was retrieved from the New Jersey School Report Cards (New Jersey Department of Education 2003, New Jersey Department of Education 2004a). These report cards provide contextual and performance data for all of New Jersey's schools and school districts. In this analysis schools were allocated to one of three groups: schools that failed to make AYP for two consecutive years, schools in the Abbott districts and schools who were neither Abbott schools nor failed to make AYP (here designated as 'other' schools). Some of the characteristics of these schools were compared with those of all New Jersey schools and appear in table 2 below.

Table 2: The characteristics of New Jersey's schools

	Schools not making AYP		Abbott schools		Other schools		All schools	
	N	%	N	%	N	%	N	%
Mean attendance rate	201	93	458	93	1755	95	2250	95
Mean school size	201	591	458	632	1747	583	2234	592
Students with LEP	201	8	458	11	1762	3	2257	5
Students speaking English at home	201	71	458	65	1762	86	2257	82
Students with disabilities	201	12	458	12	1762	13	2257	13
Mobility rate	201	24	458	24	1762	10	2257	13
Suspension rate	192	11	401	10	1246	6	1683	7
Exclusion rate	192	0.0005	401	0.002	1246	0.006	1683	0.005
Teaching experience*	104	53	235	51	977	55	1228	54
Teacher salary*	109	54	242	52	1013	57	1269	56
Teachers with temporary certificates	194	4	457	5	1275	2	1764	2.5

* percentage of teachers earning above or equal to the median state salary (or teaching experience) for faculty members in that type of school.

In several respects, few differences seem apparent between the schools not making AYP, and designated as 'failing', and the remainder of New Jersey's schools. For example, all schools report similar attendance rates, proportions of students with disabilities and teachers with comparable salaries and levels of experience. However, more often, the schools which fail to make AYP appear to be similar to the poorer Abbott schools, particularly on indicators that may be considered as proxies for poverty. For example, compared with New Jersey schools overall and the schools labelled as 'other', both Abbott schools and those that failed to make AYP have smaller proportions of students who speak English as their first language at home, but higher student mobility rates and more students with limited proficiency to speak English (LEP). Interestingly, although the student suspension rate appears to be higher in Abbott and 'failing' schools, the student exclusion rate is slightly higher in schools designated as 'other'. However, the proportion of students who are excluded from all schools in New Jersey is relatively small and slight trends like those seen here ought to be treated with caution.

The apparent relationship between whether a school is located in a wealthy or poor district and its achievement on state proficiency tests is not restricted to the characteristics outlined above, but appears to be confirmed by the distribution of schools presented in table 3 below. Three-quarters of New Jersey's failing schools are located in the least wealthy, or category A districts, while only around 12% are located in the medium to high wealth areas (categories C or above).

Table 3: Success in New Jersey's schools, according to school district wealth category.

District category	Schools not making AYP		Abbott schools		Other schools		All schools	
	N	%	N	%	N	%	N	%
A	149	74	360	78	31	2	398	17
B	28	14	81	18	171	10	263	11
CD	17	8	20	4	197	11	229	10
DE	6	3	0	0	343	19	349	15
FG	1	0.5	0	0	292	16	293	13
GH	0	0	0	0	306	17	306	13
I	0	0	0	0	405	23	405	18
J	0	0	0	0	35	2	35	1

Source: New Jersey Department of Education 2004a.

Another concern raised by some critics of NCLB is that the strict rules for student subgroups would mean that students from these groups would be over-represented in poorer or failing schools. These concerns are considered below. However, the New Jersey school report card, from which much of the data used in this analysis was retrieved, provides little detail on the demographic make-up of New Jersey's schools. In order to obtain some estimation of the proportion of students who may be from the various student subgroups, it is necessary to use data derived from the school's reporting of the numbers of students from each subgroup who participated in state-wide tests, such as the Grade Eight Proficiency Assessment (GEPA). Unfortunately, this is an imperfect measure as it only tells us about the composition of the students in that grade, rather than about the school as a whole. In addition, for reasons of confidentiality, the New Jersey

school report card datasets suppress student numbers in cases where cell sizes are lower than ten. This means that for a cohort of 160 students, if fewer than ten students were from a Hispanic background, the numbers of Hispanic students would not be displayed. Where this happens, it is often possible to calculate the values for these suppressed cells by taking into account the numbers of students in the other student sub-groups, where this is not the case the numbers of students had to be estimated. Table 4 presents the percentage of students from the main student subgroups who participated in the GEPA literacy tests in 2003.

Table 4: Mean percentage of students from the main subgroups participating in GEPA(L)

Percentage tested who were…	Schools not making AYP		Abbott schools		Other schools		All New Jersey schools	
	N	%	N	%	N	%	N	%
African-American	192	57	330	46	1297	9	1664	17
White	192	10	331	12	1298	72	1666	59
Hispanic	192	29	330	37	1297	9	1664	15
Asian	192	2	332	2	1297	7	1666	6
General Education	193	78	333	76	1306	83	1676	82
SEN	193	13	333	14	1306	14	1676	14
LEP	193	7	333	8	1306	3	1676	4
Economically disadvantaged	193	74	333	75	1306	15	1676	28

Note: the total number of schools differs from those in the tables above for two reasons, in some schools small cell sizes are suppressed and had to be excluded, also earlier tables include schools with younger students who did not sit GEPA

Schools categorised here as ‘other’, that is schools which were neither designated as Abbott nor schools that failed to make AYP, tested lower proportions of students from the African-American and Hispanic communities and higher proportions of students from the white community. These differences can, in fact be quite staggering. For example, in all of New Jersey’s public schools, while around 17% of students who were assessed using the GEPA were from the African-American community, almost 60% of these students were being taught in schools that were failing to meet state standards, according to NCLB. On the other hand only around 10% of students being assessed in these

'failing' schools were white. Similar inequities can be found in the distribution of students from economically disadvantaged homes, who again were concentrated in schools that failed to demonstrate Adequate Yearly Progress towards full proficiency. In New Jersey, students who do not follow special education programmes will follow the 'general education' route. The proportions of these students are lower in the Abbott schools and in those that failed to make AYP.

Testing higher proportions of certain student sub-groups in 'failing' schools is only really an issue if these students actually achieve lower results. Using the GEPA literacy test as an example, table 5 shows the proportions of students from the various subgroups who achieved or surpassed proficiency levels in 2003.

Table 5: Mean percentage of students making AYP in the GEPA literacy test

Student sub-group	Number of schools	Students making AYP (%)
White	514	83
African-American	283	52
Hispanic	325	60
SEN	429	31
LEP	79	16
Economically disadvantaged	406	53
All students	655	73

Source: New Jersey Department of Education 2004a.

Students from the African-American and Hispanic communities were less likely to reach minimum proficiency levels on the GEPA literacy assessment than students from white families. Students with LEP, SEN and from economically disadvantaged homes also had relatively low success rates on this test. It is also apparent from table 6, that these less successful students are concentrated in the least wealthy school districts. For example, in the 159 district A schools, less than 10% of students came from a white background. This contrasts with schools in more wealthy districts where the school population is overwhelmingly white.

Table 6: Distribution of selected student subgroups in different wealth districts.

Mean percentage who were…	…African-American		…white		…economically disadvantaged	
	N	%	N	%	N	%
A	158	46	159	9	159	79
B	84	15	84	53	85	35
CD	63	13	63	69	65	23
DE	90	8	90	76	90	13
FG	76	7	76	76	76	10
GH	72	8	72	78	73	7
I	97	4	97	80	97	3
J	12	1	12	87	12	1
All	652	18	653	57	657	30

Source: New Jersey Department of Education 2004a. N=number of schools

So on the one hand, New Jersey has established a seemingly equitable school system, particularly where control and change can make an impact at the institutional level, such as ensuring higher levels of funding for schools in less wealthy districts and similar distributions of experienced and salaried teachers. However, differences are revealed in the demographic make-up of New Jersey's schools which appear to confirm some of the fears of commentators who claim that NCLB, rather than making all schools accountable for student progress regardless of the composition of their intake, has resulted in schools in poorer districts with large proportions of students from minority and disadvantaged communities being labelled as failures. This absence of value-added measures of school success which take account of the prior attainment of students as well as their background characteristics, is important. We know from school effectiveness research both in the USA and the UK that the school accounts for a relatively small proportion of the variation in school academic outcomes, typically 8-20% and this includes error components, with by far the largest variation in outcome being attributed to student background characteristics (Jencks 1972, Reynolds 1994, Sammons *et al* 1995). Relying solely on uncontextualised raw test scores as a means of allocating success and failure to schools, fails to take account of important differences between the types of students who attend these schools, this, in turn, may result in otherwise successful schools being unfairly labelled as failing.

The following section takes these background characteristics into account when developing a value-added assessment of school achievement in New Jersey which will allow us to look more closely at schools that fail to meet state standards and consider the evidence for their apparent underachievement.

New Jersey's failing schools?

The dataset for all the schools that made Adequate Yearly Progress during the first two years of *No Child Left Behind* in New Jersey was used to create a model which, by taking into account both school and student characteristics, was used predict performance in each of the state's proficiency assessments. This model was then used to predict the test outcomes for schools that did not make AYP in this period. For brevity, only the model relating to achievement in the Grade Eight Proficiency Assessment in literacy (GEPAL) is described here, although the models produced for the other state assessments revealed similar patterns. The most powerful model related the proportion of students who achieved proficiency levels or higher in the GEPA(L) to a range of variables which may be considered to be proxies for poverty, such as the proportion of students tested who come from economically disadvantaged homes, and the relative distribution of local sources of funding for schools. The model coefficients for the multiple linear regression analysis are given in table 7 below.

Table 7: Coefficients for achievement in GEPA literacy assessments.

Variable	Coefficient
Constant	-148.37
% economically disadvantaged tested	-0.23
Attendance rate	2.46
Student mobility	-0.24
Disability rate	-0.22
Funding from local sources	0.08
% white students tested	0.02
% students with SEN tested	-0.17
Number of schools = 543	R^2 = 0.72

Around 72% of the variance in test scores can be explained by the variables listed above. Using this model, it is possible to calculate the predicted scores for the 112 schools that failed to make AYP but tested students using the GEPA(L) and, by comparing the difference between their predicted and actual scores, to consider the extent to which these schools are underachieving. In this study, schools were designated as underachieving if the z-score of the difference between their predicted and actual scores was less than –1.

Consider, for example school A which had failed to make Adequate Yearly Progress for two consecutive years and whose parents were being given the opportunity to move their children to other schools in that district. In this school, around 52% of students achieved proficiency levels or higher in the GEPA(L), while our model predicted that 40% of students would achieve this level. This means that, on this model, students in this school actually did better than expected, when the variables listed in table 7 were taken into account. Therefore, on this measure, there is no evidence to suggest that this school was underachieving. In fact, 53% of the 112 schools which, on raw score measures were labelled as failing under *No Child Left Behind* were, when background characteristics are taken into account, actually achieving higher sores than expected and therefore incorrectly labelled as failing. On the other hand, consider school B. This school is located in a relatively wealthy FG category district. The school was making AYP and, with 62% of students achieving proficiency levels or above on GEPA(L), was not subject to any of NCLB sanctions. However, on our value-added model, over 78% of students should have been working at or above proficiency levels – so on this measure, this seemingly successful school was actually 'underachieving'.

Discussion

Analysis of early results from state-wide assessments that are linked to the *No Child Left Behind* accountability requirements suggest, that in New Jersey at least, some of the concerns voiced by critics of the legislation appear to be well founded. Although

equitable in its intent, the failure of the Act to provide contextualised or value added analysis of assessment data means that many schools appear to be unfairly labelled as failing. Schools that are failing to demonstrate Adequate Yearly Progress are overwhelmingly those that are located in the poorer school districts and who serve disproportionately larger numbers of students who traditionally do less well in school, such as students from economically disadvantaged homes and who come from the African-American community. On the other hand, schools located in the more wealthy school districts were more likely to serve larger communities of white students who were performing at higher levels on the state assessments. These schools were more likely to meet New Jersey's accountability targets. The use of a value added model to account for the failure of schools to make AYP revealed that around 50% of schools were incorrectly labelled as failing and, it could be argued, being unfairly subjected to the punitive sanctions that are administered to schools that fail to make AYP for two consecutive years.

However, it has to be remembered that these are still early days. The need to take a longitudinal perspective on the development of this Act is crucial. If the legislation is actually seen to make an impact on the achievement of the lowest achieving groups of students then perhaps its dissenters will be encouraged to think again. Even so, how *No Child Left Behind* will continue to work in practice does remain to be seen. Some commentators foresee a shift from the current relatively coercive accountability measures that underpin the first manifestations of the Act, to 'softer' measures in which the consequences of failing to meet annual accountability targets will be less severe (Hess 2003, Hanushek 2003, West and Peterson 2003). Recent concessions over the assessment of students with SEN and LEP, the designation of highly qualified teacher certification and the minimum numbers of students required for test participation might suggest a softening of accountability measures (Popham 2004a, Department of Education 2004a). The use of the Act's 'safe-harbour' provisions may also be developed to allow increased flexibility in demonstrating AYP, particularly for some student sub-groups. But at least in

the short-term, it is through judicial means and pressure on legislatures that many school districts hope to see the strict accountability rules relaxed (for example, Almond 2004). However, one thing that does appear to be certain, is that now that George W Bush has secured a second term of office, *No Child Left Behind*, is here to stay.

References

Abedi, J., (2004) The No Child Left Behind Act and English Language Learners: Assessment and Accountability Issues, *Educational Researcher*, Vol. 33, No. 1, pp4-14.

Adams, K.L., Adams D.E., (2003), *Urban Education: A reference handbook*, California: ABC-Clio Inc.

Almond, A., (2004), *Educators predict lawsuits challenging No Child Left Behind*, Associated Press, 3rd November 2004, www.boston.com/news/nation/articles, accessed November 2004.

Anyon, J., (1997), *Ghetto Schooling: A Political Economy of Urban Educational Reform*, New York: Teachers College Press.

Bylsma, P.J., (2004), '*Challenges and successes in Washington*', paper presented at the American Educational Research Association Annual Meeting, San Diego, April 12-16, 2004.

Chester, M.D., (2004), '*Ohio's experience with AYP implementation, Fewer Schools than Expected*', paper presented at the American Educational Research Association Annual Meeting, San Diego, April 12-16, 2004.

Department of Education (2002), No Child Left Behind: Executive Summary, www.ed.gov/nclb/overview/intro/presidentplan/page_pg3.html, accessed June 2004.

Department of Education, (2002a), Education Statistics Quarterly, US Department of Education, http://nces.ed.gov/Pubs2003/Hispanics, accessed May 2004.

Department of Education (2004a), Secretary Paige announces new policies to help English language learners, Press Release 19 February 2004, www.ed.gov/news/pressreleases/2004/02/02192004.html, accessed August 2004.

Education Week (2004a), Accountability conflicts vex schools, *Education Week*, 10th March 2004, Vol. 23, No. 26, p1.

Education Week (2004b), Postcard from Independence, Mo. *Education Week*, 9th June 2004, Vol. 23, No. 39, p1.

The Education Trust, (2002) *The Funding Gap: Low-income and Minority Students Receive Fewer Dollars*, The Education Trust, August 2002, www2.edtrust.org/EdTrust/Product+Catalog/special+reports.htm#2003 [accessed May 2004].

Firestone, W.A., Goertz, M.E., Natriello, G., (1997), *From Cashbox to Classroom: The struggle to fiscal reform and educational change in New Jersey*, New York: Teachers College Press.

Flanagan, A., Grissmer, D., (2003), The Role of Federal Resources in Closing the Achievement Gap, in *Bridging the Acheivement Gap*, Chubb, J.E., Loveless, T., (eds.), Washington DC: Brookings Institute.

Hochschild, J., (2003), Rethinking Accountability Politics, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds.), Washington DC: The Brookings Institution.

Hess, F.M., (2003), Refining or Retreating? High-Stakes Accountability in the States, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds.), Washington DC: The Brookings Institution.

Haney, W., (2000), The Myth of the Texas Miracle in Education, *Education Policy Analysis Archives, 8*(41), http://epaa.asu.edu/epaa/v8n41/ accessed August 2004.

Hanushek, E.A., Raymond, M.E., (2003), Lessons about the Design of State Accountability Systems, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds.), Washington DC: The Brookings Institution.

Jencks, C., (1972), *Inequality: a Reassessment of the Effect of Family and Schooling in America*, Basic Books, New York.

Jencks, C., Phillips, M., (1998), *The Black-White Test Score Gap*, Washington DC: The Brookings Institute.

Kane, T.J., Staiger, D.O., (2003) Unintended consequences of racial subgroup rules, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds), Washington DC: The Brookings Institution.

Lee J., (2002), Racial and Ethnic Achievement Gap Trends: Reversing the Progress Toward Equity? *Educational Researcher*, 31(1), p3-12.?? Or AERA ref?

Linn, R.L., (2003) Accountability: Responsibility and Reasonable Expectations, *Educational Researcher*, Vol. 32, No. 7, pp3-13.

McNeil, L.M., (2000), Creating New Inequalities: Contradictions of Reform, *Phi Delta Kappan*, 81, pp728-735.

National Centre for Education Statistics (2002), *Digest of Educational Statistics 2002*, http:/nces.ed.gov/programs/digest/d02/tables/dt020.asp, [accessed May 2004].

National Commission on Excellence in Education (NCEE) (1983), *A Nation at Risk*, US Department of Education, Washington DC: US Printing Office.

New Jersey Department of Education (2003), 2002 State Assessment Data Analysis for Schools In Need of Improvement, http://www.state.nj.us/njded/data/title1/, accessed November 2004.

New Jersey Department of Education (2004), *State of New Jersey, Consolidated State Application Accountability Workbook: No Child Left Behind in New Jersey*, http://www.state.nj.us/njded/grants/nclb/workbook.pdf, accessed November 2004.

New Jersey Department of Education (2004a), New Jersey's Report Cards, Meeting the Requirements of State Law and the *No Child Left Behind Act*, accessed November 2004, http://education.state.nj.us/rc.

New Jersey Department of Education (2004b) NJ Department of Education District Factor Groups (DFG) for School Districts, accessed November 2004, http://www.state.nj.us/njded/finance/sf/dfgdesc.shtml.

Orfield, G., (2000), Policy and Equity: Lessons of a Third of a Century of Educational reforms in the United States, in *Unequal Schools, Unequal Chances: The Challenges to Equal Opportunity in the Americas*, Reimers, F., (ed), Massachusetts: Harvard University Press.

Perry, R., (2004), *Implementing AYP in an existing accountability system in California*, paper presented at the American Educational Research Association Annual Meeting, San Diego, April 12-16, 2004.

Popham, W.J., (2004), *America's 'Failing' Schools, How Parents and Teachers can cope with No Child Left Behind*, New York: RoutledgeFalmer.

Popham, W.J., (2004a) Shaping up the 'No Child' Act: Is edge-softening really enough? *Education Week*, Vol. 23, No. 38, p40, 26th May 2004.

Ravitch, D., (1995), *National Standards in American Education*, Washington DC: The Brookings Institution.

Reynolds, D., (1994), School Effectiveness and Quality in Education, in *Improving Education: promoting quality in schools*, Ribbins, P. and Burridge, E., (eds.), Cassell, London.

Rudalevige, A., (2003) No Child Left Behind: Forging a Congressional Compromise, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds.), Washington DC: The Brookings Institution.

Sammons, P., Hillman, J., Mortimore, P., (1995), *Key Characteristics of Effective Schools – a review of school effectiveness research*, OFSTED, London.

Schmidt, W.H., McKnight, C.C., Cogan, L.S., Jakwerth, P.M., Houang, R.T., (1999) *Facing the Consequences. Using TIMSS for a Closer Look at US Mathematics and Science Education*, Dordrecht: Kulwer Academic Publishers.

West, M., Peterson P., (2003), The Politics and Practice of School Accountability, in *No Child Left Behind? The Politics and Practice of School Accountability*, Peterson, P., and West M., (eds.), Washington DC: The Brookings Institution.

Biographical note and acknowledgements

Dr Emma Smith is a Research Fellow in the Department of Educational Studies at York University. Her research interests include underachievement and differential attainment in secondary schools, educational inequalities and assessment.

The author would like to thank Professor Bill Firestone of Rutgers University, New Jersey and Professor Stephen Gorard of York University for their ideas and comments on this paper.