Education Policy Analysis
Archives/Archivos Analíticos de Políticas
Educativas
ISSN: 1068-2341
epaa@alperin.ca
Arizona State University
Estados Unidos

Gregory Harman, Wm.; Boden, Camille; Karpenski, Jeremy; Muchowicz, Nicole
No Child Left Behind: A Postmortem for Illinois
Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 24, 2016, pp. 1-24
Arizona State University
Arizona, Estados Unidos

education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal

epaa aape

Arizona State University

Volume 24 Number 48 April 18, 2016 ISSN 1068-2341

No Child Left Behind: A Postmortem for Illinois

Wm. Gregory Harman
Camille Boden
Jeremy Karpenski
&
Nicole Muchowicz
Dominican University, Chicago
United States

Citation: Harman, W. G., Boden, C., Karpenski, J., & Muchowicz, N. (2016). No Child Left Behind: A postmortem for Illinois. *Education Policy Analysis Archives, 24*(48). http://dx.doi.org/10.14507/epaa.v24.2186

Abstract: In this study, the outcomes of No Child Left Behind (NCLB), as implemented in Illinois, are evaluated in terms of high school standards testing results between 2003-2013. NCLB was a policy dedicated to closing the gap in schooling outcomes nationally in the space of a decade. There have been few systematic examinations of its macro-level results for those exiting high school, especially considering the attention, time, effort, and money dedicated to its implementation. It has been subsumed into newer reform policies that move forward from the same assumptions and structures without a look back. This is a macro study of the outcomes in one state, Illinois, using its assessment system. Data include Prairie State Achievement Examination (PSAE) results in reading and math as well as graduation rates from high schools. The data is evaluated across the

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 15/7/2015
Revisions received: 22/12/2015
Accepted: 23/12/2015

state as a whole and within categories of urban, suburban, town, and rural. Outcomes in reading, math, and graduation rates remain unchanged across the decade at the state and all community-type categories. Potential problems with implementation and design of NCLB are proposed with the intention of informing current and future policy, especially in regard to continuing a standards/accountability regime under the Common Core.
Keywords: No Child Left Behind; NCLB; standards; testing; standardized tests; Illinois; Chicago; Common Core; policy; performance; compliance; Prairie State Achievement Examination; PSAE; quantitative research

No Child Left Behind: Una autopsia para Illinois

Resumen: En este artículo, los resultados de la ley *No Child Left Behind* (NCLB), en Illinois, son evaluados en cuanto a los resultados de las pruebas de secundario entre 2003-2013. NCLB fue una política enfocada a cerrar la brecha nacional en los resultados escolares en una década. Hubo pocas examinaciones sistemáticas de los resultados al nivel macro para los que terminaban la escolaridad, especialmente considerando la atención, el tiempo, esfuerzo y dinero dedicado a la implementación de NCLB. Esta reforma ha sido subsumida en nuevas políticas de reforma con las mismas asunciones e estructuras sin considerar su pasado. Este es un estudio macro de los resultados en Illinois, utilizando su sistema de evaluación. Incluimos los resultados en lectura, matemáticas y las tasas de graduación de escuelas secundarias del sistema de examinación Prairie State Achievement Los datos se evalúan para el estado entero y usando las categorías de urbano, suburbano, pueblo y rural. Los resultados en lectura, matemática y tasas de graduación se mantuvieron sin cambios en todas las categorías en el estado y tipo de comunidad en la década evaluada. Se identifican problemas con la implementación y diseño de NCLB con la intención de la política actual y futura, especialmente con respecto a continuar un régimen de estándares/responsabilidad bajo el Programa Common Core.
Palabras-clave: No Child Left Behind; NCLB; estandares; exámenes; pruebas estandarizadas; Illinois; Programa Common Core (CCSS); desempeño; Prairie State Achievement Examination; PSAE

No Child Left Behind: Uma Autopsia para Illinois

Resumo: Neste artigo, os resultados da legislação No Child Left Behind (NCLB) na sua implementação em Illinois, são avaliados em quanto aos resultados da prova da escola secundaria entre 2003-2013. NCLB foi uma política dedicada a fechar a brecha nos resultados escolares a nível nacional no lapso de uma década. Nesse período se realizaram poucas analises sistemáticas dos resultados ao nível macro para os que concluíram a escola, especialmente em consideração da atenção, tempo, esforça y dinheiro dedicados a sua implementação. NCLB foi subsumida em novas políticas de reforma que utilizam as mesmas premissas e estruturas sem considerar os antecedentes. Este es um estudo macro dos resultados em Illinois, utilizando seu sistema de avaliação. Incluídos nos dados são os resultados em leitura y matemática, os índices de formação das escolas secundarias do sistema de avaliação do Prairie State Achievement Examination. Os dados são avaliados por o estado inteiro y dentro de categorias de urbano, suburbano, cidade y rural. Os resultados em leitura, matemática e os índices de formação não mudaram em todas as categorias no estado e tipo de comunidade através da década. Analisamos os problemas com a implementação y desenho de NCLB com o obejtivo de informar a política atual y futura,

especialmente com respeito a continuar um regímen de exames e sistema de consequências severas no Programa Common Core.
Palavras-chave: No Child Left Behind (NCLB); exames; testes padronizados; Chicago; Programa Common Core; CCSS; politica; Prairie State Achievement Examination (PSAE); pesquisa qualitativa

No Child Left Behind: A Postmortem for Illinois

No Child Left Behind, initiated in 2003, was the first national attempt to apply the logic of standards-based schooling. The logic model of the standards movement is: if students display mastery of high quality standards, then they have received a quality education. If teachers are held accountable for transmitting the high quality standards, then students will gain mastery of those standards. If high-stakes tests are aligned to the standards, then teachers and schools will be held accountable to having transmitted the standards. Thus, creating standards and then tests to assess student mastery of them, and then applying those tests such that failure to meet mastery targets (Adequate Yearly Progress, AYP) results in drastic, negative consequences for schools and teachers will result in quality education for all students. To meet the promise of this model, by 2014,100% of students in public schools across the nation were to be passing their states' standardized tests according to the No Child Left Behind law. The NCLB goal was to have all schools and all students demonstrate proficient performance in reading and mathematics by 2013-14. This level of success eluded the states, and expecting such a level of success in the first place was unrealistic. Yet there has been little public or scholarly discussion in the aftermath of successes and failures of No Child Left Behind as policy or as a useful model for the alleged ultimate beneficiaries: students completing high school. The largest quantity of research regarding NCLB occurred during application, was targeted at interim progress by examining fourth and eighth grade outcomes, and focused upon issues of implementation and to what extent testing gaps were being closed. There has been little attempt to garner lessons in retrospect about the overall outcomes of the policy. The nation has moved directly into controversies about applying the Common Core, attempting to move forward in essentially the same direction without looking back. Yet it is likely that the record of No Child Left Behind in the states leaves us with information that can help us guide the direction of our next steps in public schooling.

The research question was: what improvement did students finishing high school across the state of Illinois make according to the state's standardized tests? This research employs the eleventh-grade high school tests as the essential measure for all students in the state. The intent of No Child Left Behind was that all students would graduate high school with passing scores indicating proficiency in reading and math, with both the tests and passing scores indicating proficiency established by their states. Therefore, an essential question is how well the state progressed towards this goal. Answers suggest directions for further questions and decisions in public school policy. The process of exploring the data is revealing in regard to the collection of data across an entire, diverse state, home to one of the nations' largest cities with one of the most researched urban school systems, but also to countywide systems, suburban districts, and rural districts.

Literature

Most studies of NCLB testing outcomes concern fourth and/or eighth grades. These studies are only tangentially of interest for this research. Overall, the results of these studies relate to this study in that they also show a lack of systematic progress in test outcomes or narrowing of testing gaps in Illinois (see Council of Chief State School Officers, 2011, pp. 8, 10, 14, 16; National Center for Education Statistics, 2013, pp. 103,105; Rosaen, Schwartz, and Forbes, 2007, in Ballou 2008; Rouse et al., 2007).

One study of fourth and eighth grade outcomes is worth mention because it features some of the most considered quantitative methodology in its nuance and approach to the complex situation of comparisons between states and over the timespan of NCLB. This is a study by Lee and Reeves (2012) that employed a "comparative interrupted time series analysis" of National Assessment of Educational Progress (NAEP) results that uses "hierarchical linear modeling latent variable regression with inverse probability of treatment weighting" to account for the discontinuities between state NCLB adoption patterns, performance growth, and level changes (pp. 209, 213). It uses elegant means to account for many of the shortcomings of previous studies. Even with such involved measures, the authors come to a similar conclusion to most other outcome studies (and to this one that purposefully employs far less sophisticated metrics): Changes in performance have been minimal, with slight decreases in achievement gaps in math, but increases in the achievement gap in reading nationwide in both the fourth and eighth grades (p. 224).

The pattern of slight improvement in mathematics but static performance in reading at the middle school level is seen again in a study sponsored by the Council of Great City Schools (2010), specifically focusing on Chicago. They report small progress, and only in math, using (NAEP) as the measure. It is worth mention since it focuses on Chicago, where the majority of the students reported in this study reside. In the city, there has been progress in the eighth grade in math with a 9% drop in the percentage of students testing below the basic level and a 6 % increase in the students scoring at or above the proficient level (p. 40).

Other comparative outcome data that includes Illinois comes from Bandeira and colleagues, reporting for the U.S. Department of Education. In 2005 and 2007, on NAEP assessments, Illinois scored 25th in the nation in reading for the fourth grade, and 36th for the eighth grade. In math, Illinois was 41st in the nation for the fourth grade and 43rd for the eighth grade. The low place of Illinois is of interest for this study as a corollary to the lack of progress within the NCLB framework as Illinois pursued it. Peterson and Lastra (2010) believe that Illinois' low ranks stem from state proficiency levels and assessments that did not correlate well with the NAEP definitions of proficient (pp. 35−36).

All of the studies mentioned above focus upon elementary and middle school outcomes. The only analysis found for the high school level in Illinois, the particular concern of this study, came from The Consortium on Chicago School Research (2005). Their findings indicate small improvements in Chicago in both reading and math, but with statistical significance only in math (p. 10). As they conclude, "student performance on the PSAE, both in CPS and in Illinois as a whole, has remained at about the same level since the test's first administration in 2001" (p. 24). This agrees with the essential conclusion of this work.

The Illinois State Board of Education (ISBE, 2012) reported that the percentage of low-income students enrolled in Illinois schools rose between 2003 and 2012 from 37.9% of enrollment to 49.0%. (p. 3). This change could have made an impact upon outcomes and achievement gap. The reasoning is worth mention, since the logic that scores were

disappointing because of the increased participation of the very population targeted for improvement is at odds with the intent of the policy.

ISBE also reported that the average scores on the ACT in the state (upon which the state high school standards test is based), increased only 0.6% from 2003 to 2012. Similarly, the ACT Corporation's (2014) own reports of trend data of average groupings on the Prairie State Achievement Examination showed the same lack of any pattern of growth or shift as the ISBE report. This concurs with the analysis in this work.

There were other approaches to the outcomes question that relate to the discussion of the results in this work. Sims (2007) focused on the increased likelihood that the presence of at-risk subgroups would determine schools' AYP status, and thus their responsiveness to NCLB sanctions. The study found that NCLB had been counterproductive for achievement in schools with large numbers of at-risk subgroups, partially because of the measures that schools and teachers take to address the high-stakes examinations.

The notion that teaching so that students will score better on high-stakes tests became acceptable and even desirable in the decade of NCLB, at least when the test aligned with high-quality standards. Yet this scenario turns out to be no more beneficial in terms of the information or the outcomes obtained than teaching to poorly aligned tests. Koretz (2005) explains how "test-based accountability can lead to misleadingly large gains in scores" that misrepresent what the students actually have gained (pp. 5-6). The reason is that testing attempts to judge outcomes in an entire domain of knowledge based upon a small sample from that domain. In order to have students score better on particular high-stakes tests, teachers and schools learn quickly how to address key performance elements. They reallocate their resources, align their instruction and assessment to resemble the tests, and coach the students on format and testing protocols (pp. 11-14). The more successful these forms of teaching to the test, the wider the rift between test performance and the real outcomes in the entire domain. The items and the means learned specifically for the test represent a tiny portion of knowledge in the entire domain, but a very large portion of the sample from the domain that is tested (pp. 3-5).

This explains sudden drops in scores and then subsequent rises when tests are changed (as occurred in Illinois in the 2014-15 school year when the PARCC examination replaced the PSAE). Schools learn to successfully teach to a test, which makes the students seem to have beneficial outcomes. Then, when the sample and the performance elements shift significantly, the students' scores drop back to what is likely a more valid representation of their knowledge in the total domain (p. 6).

There is no specific policy nor allocation of resources to control for the various means of score inflation, nor for even more questionable practices that cause, in the words of Haladyna, Nolan, and Haas (1991), "score pollution" (p. 4). In addition to Koretz' set of inflating practices, Haladyna, Nolan, and Haas list distortive practices such as dismissing low-achieving students on testing days, the kind of practice also noted by Vasquez-Heilig and Darling-Hammond (2008). They questioned the benefit of the high stakes environment by illustrating how Texas schools responded by gaming the system, abandoning at-risk students rather than putting in the resources to improve outcomes. These results align precisely with what Amrein-Beardsley (2009) predicted would happen as incentives tempted administrators to "game the system" (p. 3). During the decade of NCLB, there is no indication that the State Board of Education had mandate or capacity to regulate such practices in Illinois.

Other research focused particularly on closing of achievement gaps. The National Council on Disability (NCD, 2008) remained positive on the application of NCLB because it

has meant significantly better attendance by students with disabilities and attention to their academic progress (pp. 12-14). However, the NCD was unable to report any highly significant trend in reducing the achievement gap between disabled and non-disabled populations, which aligns with the results of this study.

In racial achievement gaps, CEP (2010) reports that little progress has been made nationwide at reducing achievement gaps, another agreement with the macro analysis within a single state conducted here. In Illinois, the CEP states that there was "mixed progress made in narrowing achievement gaps in reading and math…[with] instances of gaps narrowing as well as widening at various grade/subject combinations" (online appendix, p. 1). Using mean scale scores, in grade 11 reading, the gap increased between 2006 and 2009 by 4.5%. In mathematics, the gap remained constant, with the changes for all groups going down by between 1 and 3 percentage points.

Reardon et. al. (2013) stated that there was "no evidence overall that NCLB contributed to a reduction in racial achievement gaps" (pp. 1-2). One of the observations in his study was particularly relevant to the discussion portion of this research:

> … in a school with few black students in tested grades, the test scores of black students would not be reported separately and the school would not be required to show adequate yearly progress for black students (black students' scores would still be included in calculations of the school's overall proficiency rate, however, though they might matter little given the small number of black students). In such a school, NCLB may create little or no incentive to focus attention on the performance of the small number of black students in the school—indeed, it may create an incentive to focus primarily on the performance of the schools' white students. (p. 5).

A study on race achievement by Barton and Coley (2008) employed NAEP data reaching back 25 years before NCLB to compare with progress since NCLB. They found that "scores jumped at each quartile between 1975 and 1990 but have not improved since, and have fallen in the bottom quartile since 1990." There is a similar rise between 1975 and 1990 and leveling out leading to decline by 2004 in mathematics (p. 4, 14). The high-stakes testing and accountability environment did not produce any tangible benefits, as other studies also predicted (Amrein and Berliner, 2002; Marchant, Paulson, and Shunk, 2006), a conclusion which this study supports

In sum, the literature suggests that there have been no consistently positive outcomes produced by NCLB. Some of this work has suggested reasons that NCLB failed to produce anything like the outcomes sought, but future directions for policy remain only vaguely suggested. Current policy is primarily an extension of the NCLB model of testing and accountability measures. This has occurred without a substantive mass of literature to suggest that the particular direction taken would be the most productive approach. The unproven assumptions of the current approach, based on selective reading of research, is that the failures of NCLB arise from lack of consistency between states and the lack of particular relation of previous standards to desirable and measurable skill outcomes. Left to discover is whether these were indeed the shortcomings, and if not, what the essential shortcomings have been.

Method

State Data

The state of Illinois collected and keeps all the data that is needed to answer the research question. For accountability purposes to meet the requirements of No Child Left Behind, all eleventh-grade students in public (and most charter) schools in the state between 2003 and 2014 took the Prairie State Achievement Examination (PSAE), a reading and math skills test developed by the ACT Corporation as the final test for meeting No Child Left Behind provisions.

The PSAE was created, maintained, and validated by the ACT Corporation, which produces the ACT test for college admissions. The reading and math portions of the PSAE are the ACT reading and math tests. Additional material, called Work-Keys, and administered separately, was designed to compensate for the fact that the ACT was designed as a test for admissions to selective universities. The ACT plus Work Keys is the Prairie State Achievement Examination. The Center on Education Policy (2009) reported that there was a concern that the test was not specifically designed to correlate with the state's standards (p. 2). However, the *Technical Manual* for the PSAE (ACT, 2013, argues in detail that the connections between the ACT component and the Illinois Standards are well thought-out and were validated by a task force including Illinois teachers (pp. 5-11). The *PSAE Technical Manual* also describes how the validity of the test in measuring the target skills was assured by the processes that ACT employs that are widely accepted for research and reporting (ACT, 2013, pp. 5, 8-29). In the researchers' own comparison between the ACT's college readiness index and the PSAE scores for sample communities, the same parallels are evident.

A positive factor in using this data is the stability reliability of the instrument. As Consortium on Chicago School Research, University of Chicago (CCSR) (2005), and the *PSAE Technical Manual* (ACT, 2013) describe, the PSAE maintained a constant level of difficulty through the decade, and the state maintained a cutoff score of 133 on the exam throughout. In terms of the skills, students passing in any of the years are being measured in an equivalent way.

Passing percentage data overall and by various demographic groupings was collected by the state and made public via the online Illinois Interactive Report Card. The passing percentage for a school is a sum of the rate of students achieving a designation of "Meets Standards" and "Exceeds Standards" on the constituent sub-tests.

There is no place that the state aggregates the data to report the overall situation of the state. The state was not forthcoming when the researchers for this project requested the database in a form that could be more readily aggregated and examined. The researchers were sent two files that could not be translated by Excel or by common statistical database software. There was no reply to requests for a more usable version of the data.

Research Method

Therefore, to answer the research question, the researchers collected the data by copying it, line-by-line, for every school for every year, from the Illinois Interactive Report Card. They transcribed into Excel the passing percentages on the Prairie State Achievement Examinations reported for all eleventh grade students from every reporting high school in the state of Illinois from 2003 to 2013 for reading and for math. Each of the four researchers was assigned to cross-check entries from the others so that correct data entry was assured. To insure that they were left with a database that could provide more

information for future research, the researchers also copied all of the pass rates by demographic categories and each high school's graduation rates for the year.

The result is a spreadsheet containing every pass rate from every reporting Illinois high school in reading and math for every year for each reported demographic ranging from 2003-2013. The pass rate reported for the school indicates the sum of percentages of students achieving "Meets Standards" and "Exceeds Standards" scores on the PSAE sub-tests. To answer the research question, a mean of the percentage pass rates for reading and for math for the state were created for each year and compared. Additionally, the researchers compared passing rates with graduation rates, for which the graduation rate was listed for the test year plus one to account for the test-takers' own graduation rates (juniors take the test, and of course most often graduate a year later). A Runs test was used through SPSS to determine the significance of variance across years for aggregated data.

The Wald-Wolfowitz Runs test (RUNS in SPSS Statistics software) is a non-parametric statistical test used to determine if runs in a data set are mutually independent. It can be used to determine whether or not there is a pattern when there is only one variable over a set of trials or over discrete, repeated data over time. This is done by assigning positive and negative values to figures above and below the median respectively to establish runs. It reveals progress or regress, even when such movement might not be immediately obvious, and provides a measure for significance. The null hypothesis is assumed, with absolute values greater than 1.96 generally indicating significance at the 5% level.

The method is designed purposefully as a blunt metric. It keeps the exploration at a macro level, comparing the final results for the entire state population that was supposed to receive the ultimate benefits of the policy: all students exiting high school. More sophisticated and fine-grained analyses are quite revealing, but tend to obscure rather than reveal the most vital true story of the outcome. Moreover, this method produces information highly accessible to citizens and policymakers, which the entire national effort was premised upon producing. The more sophisticated the analysis, the less comprehensible it becomes to policymakers and citizens. It quickly gets to a place where the information derived from the data is only accessible to quantitative policy scholars. The possibility to have a positive impact upon policy diminishes. The method employed in this study is used because it is summative, concise, yet still accurate and truthful in the story it tells. It has the additional benefit that it can be easily employed in other states to illuminate the patterns NCLB has produced there. Replication in every state could, in a very reasonable time horizon, create an accessible picture of NCLB's national effects, and have an impact upon future policy decisions.

Data

Problems with the Data Set

Because this is an analysis of policy outcomes, it is worthwhile to relate that the researchers encountered issues during transcription that are almost certainly inaccuracies in the state data.

There were many schools that lacked any data for early years, and then provided data for later ones. For example, schools would have no data between 2003 and 2007, then have data available from 2008-2013. This was easily explained in the case of charter schools that opened their doors during the decade, but this does not explain all the cases.

There were no schools that showed the opposite trend: schools that provided data during the beginning of the decade and lacked data towards the end. This led the researchers

to explore what happened to the high schools that were closed for poor performance. For example, Englewood High School closed its doors in 2008 due to poor performance. When the researchers searched for Englewood High School on the interactive report card, there were no entries for 2003-2007 when the school was open. Students at these schools were therefore unaccounted for.

There were also inconsistencies with data reporting. There is a particular concentration of lack of reporting by schools in 2006-2007, with no explanation available. There were many schools that reported exactly the same pass rates in both reading and mathematics, some for more than one year at a time. There were 307 records (4.51% of the total) in which the reading score was the same as the math score. Of those, twenty-eight schools reported matching reading and math scores for two separate years. Of those, seven reported matching reading and math scores in two consecutive years, of which two reported exactly the same scores for both of those two years. Five schools reported matching reading and math scores in three years, of which one reported matching scores in two consecutive years, and one reported matching reading and math scores for three consecutive years. One school even reported matching reading and math scores over four separate years. While the figures may be accurate, the likelihood of such identical outcomes on two separate measures, especially over more than one year, is remote.

Finally, it is worth noting that there is a significant hole in the data that is an unintended consequence of policy. This is the problem described by Reardon and colleagues (2013, see quote above in literature review) that a large number of students are not reported within their demographic subgroups. NCLB required states to set the minimum number of students that constitute a subgroup for reporting. Because of privacy concerns, Illinois set their number at 40 students. If there were under 40 students in a subgroup in a school, the school did not report those students' scores as part of a subgroup. There were many schools whose percentage passage rate for white students was higher than the percentage rates for all students, but no data was reported for any other subgroup because of this cutoff.

Data Produced

The mean percentage pass rates for reading and for mathematics for the entire state are presented in Table 1 and in Figures 1 and 2. The RUNS test on the data verifies the null hypothesis for both reading (.210) and for mathematics (.682), indicating that there is no significant pattern of movement in passing percentages. The mean percentage graduation rates for the state are also presented in Table 1.

Mean percentage pass rates for reading and mathematics by community type were also calculated, and are presented in Table 2 and in Figures 3-7. RUNS tests on each community type (reading and math respectively: large cities .210, .540; other cities, .210, .210; rural, 1.0, .977; suburbs, .540, .540; towns, .540, .540) verify the null hypothesis for reading and math for every community type: There is no significant pattern of change. The mean percentage graduation rates for community types are also presented in Table 2.

The researchers also wished to report on the number of students who, as a result of the cutoff for reporting in demographic subgroups, are missing from the record as a separate group. The estimate was created by selecting a representative sample of 100 of the schools over two years (2012 and 2013), using the total eleventh-grade enrollment for a year and the percentage of minority students for those schools, then multiplied to account for the total number of schools in the population. The resulting figure is 18,500 students.

Results

There was no significant change in the passing rates on the Prairie State Achievement Examination for either the reading or the mathematics tests between 2003-2013 (See Table 1 and Figures 1 and 2).

Table 1
Average of Passing (Reading and Math PSAE) and Graduation Rates across Illinois by Year with Standard Deviations

	Average			Standard Deviation		
	Reading All	Math All	Grad Rate	Reading All	Math All	Grad Rate
2003	54.5	50.4	88.4	15.1	17.7	10.8
2004	55.3	50.4	88.9	15.2	17.7	10.6
2005	59.1	50.6	89.1	14.3	17.6	9.8
2006	57.5	52.3	89.5	15.7	18.1	9.8
2007	53.6	51.0	88.7	16.0	18.5	11.6
2008	52.7	51.6	88.6	17.9	19.2	11.9
2009	56.6	49.9	89.0	17.8	19.9	11.3
2010	53.5	50.4	89.6	17.9	20.0	10.8
2011	51.7	51.2	84.4	18.4	20.1	9.9
2012	51.1	50.5	84.6	18.3	20.4	10.7
2013	53.0	49.1	88.6	18.1	19.8	9.2
Decade Avg.	54.4	50.7	88.1			
Decade STD	2.5	0.9	1.8			
RUNS (SPSS)	0.210	0.682				

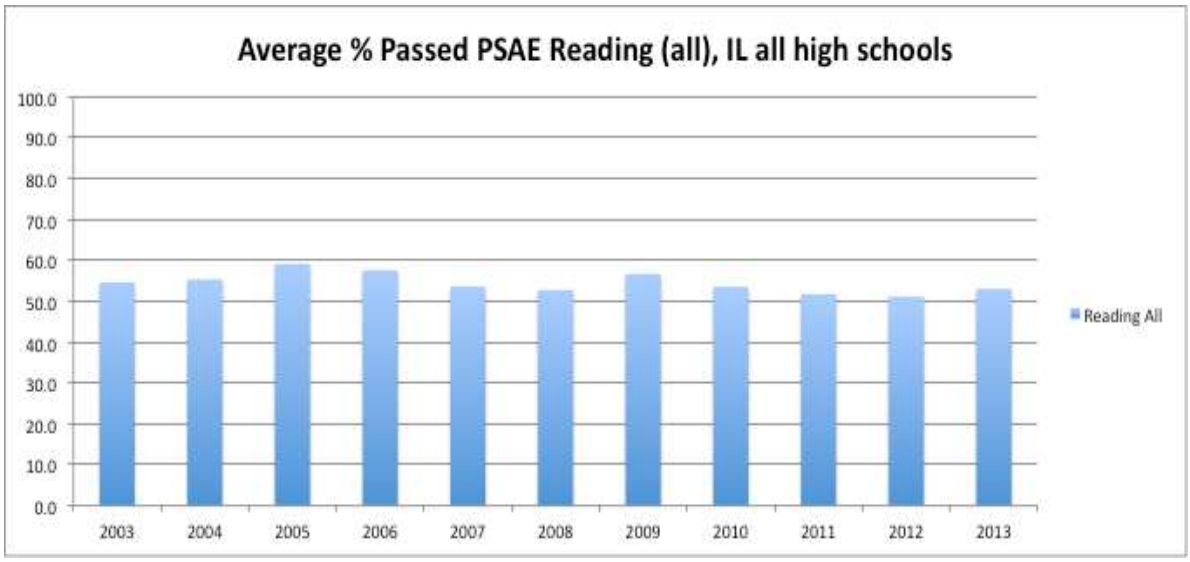


Figure 1. Average % Passed PSAE Reading, IL high schools

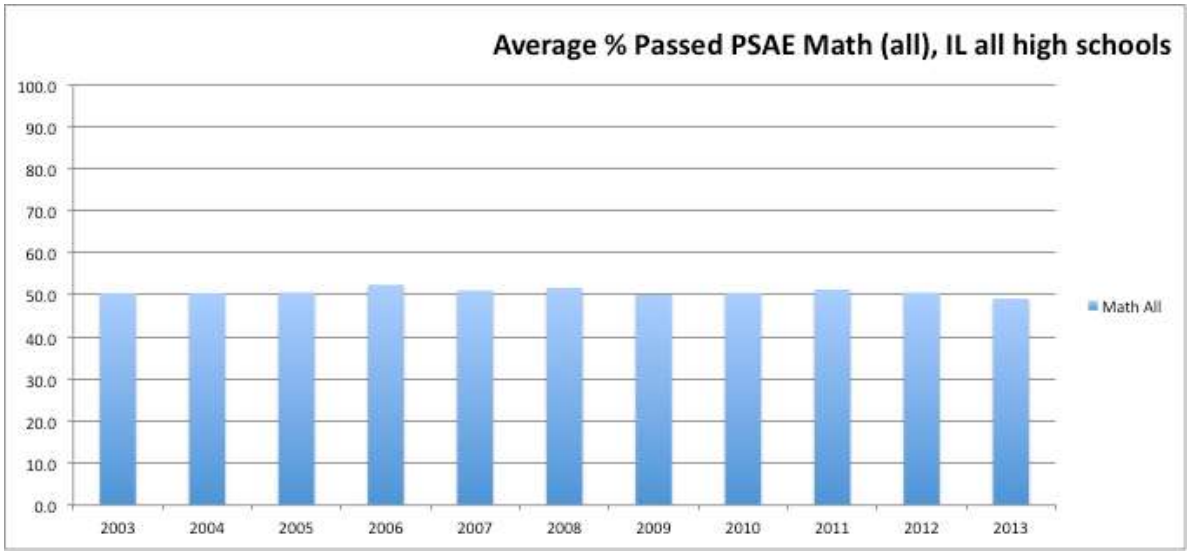


Figure 2. Average % Passed PSAE Math, IL high schools

The average pass rates across all the years for the state were 54.4% (standard deviation of 2.4) for reading, and 50.7% (standard deviation of .9) in math. In the state as a whole, students performed similarly in every year of the decade from 2003-2013. This was also true when the data was disaggregated by community type (See Table 2 and Figures 3-7).

Table 2

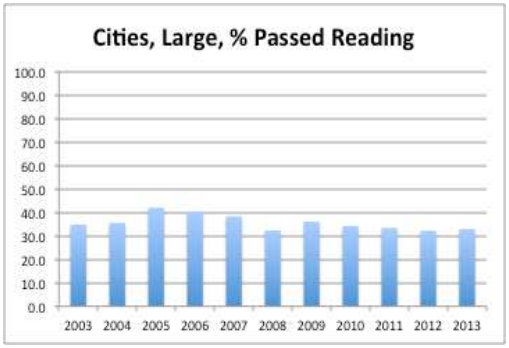
Average of Passing (Reading and Math PSAE) and Graduation Rates Across Illinois by Year with Standard Deviations by Community Type

		Reading		Math		Grad Rate	
		Avg % pass	StdD.P	Avg % pass	StdD.P	%	StdD.P
Cities, Large	2003	34.7	22.5	26.7	23.2	72.3	13.2
	2004	35.5	22.7	27.5	24.7	74.5	13.4
	2005	42.0	22.2	27.7	24.8	76.3	11.9
	2006	40.3	23.6	32.0	25.8	76.0	12.7
	2007	38.2	24.0	33.0	27.1	68.9	14.9
	2008	32.3	25.5	31.3	27.6	70.5	16.6
	2009	36.0	25.3	29.2	26.7	71.3	15.5
	2010	34.2	25.0	30.1	26.5	73.4	15.3
	2011	33.3	26.2	32.2	27.7	77.4	13.7
	2012	32.1	25.1	32.0	27.1	79.1	16.4
	2013	32.9	23.7	29.2	24.8	86.8	13.1
	Decade Avg.	35.6	24.2	30.1	26.0	75.1	14.3
	RUNS (SPSS):	0.210		0.540			
Cities, Other	2003	55.3	13.1	51.1	16.9	86.9	8.8
	2004	56.3	12.7	51.5	16.1	85.3	10.3
	2005	58.3	12.8	50.0	17.4	85.8	9.5
	2006	57.4	15.3	51.4	18.1	85.4	9.9
	2007	54.0	15.2	50.7	17.5	86.3	8.7
	2008	52.7	15.3	52.0	17.5	86.8	9.2
	2009	55.8	14.8	50.0	18.9	86.8	9.4
	2010	53.3	14.0	51.6	16.7	89.2	6.8
	2011	49.4	15.2	49.3	18.2	81.9	9.0
	2012	49.1	13.9	48.4	18.1	79.9	11.0
	2013	51.1	14.8	48.2	17.7	81.0	10.6
	Decade Avg.	53.9	14.3	50.4	17.6	85.0	9.4
	RUNS (SPSS)	0.210		0.210			
Rural	2003	57.0	10.9	52.5	13.0	91.7	8.7
	2004	58.0	10.7	52.5	13.3	92.4	8.0
	2005	61.7	10.4	53.2	12.0	92.1	7.7
	2006	60.3	12.1	55.3	14.1	92.7	7.4
	2007	56.0	12.1	53.0	14.3	93.0	7.1
	2008	56.4	12.6	54.5	14.3	92.6	7.3
	2009	60.6	12.7	53.0	14.9	93.1	6.7
	2010	57.8	12.8	53.2	15.6	93.4	6.7
	2011	56.0	13.0	55.4	14.4	87.3	7.6
	2012	56.4	12.3	54.5	15.0	87.2	7.8
	2013	57.2	11.5	53.0	13.5	91.2	5.7

Table 2, cont'd

Average of Passing (Reading and Math PSAE) and Graduation Rates Across Illinois by Year with Standard Deviations by Community Type

	Decade Avg.	57.9	11.9	53.7	14.0	91.5	7.3
	RUNS (SPSS)	1.000		0.977			
Suburbs	2003	58.4	13.6	56.3	16.7	90.2	8.4
	2004	58.5	14.0	55.7	16.5	90.6	8.7
	2005	62.1	13.0	55.7	17.5	90.4	8.4
	2006	62.2	13.9	57.3	16.8	91.0	7.2
	2007	57.8	15.2	56.7	17.4	90.4	8.2
	2008	58.1	15.9	58.6	17.2	90.8	9.3
	2009	61.5	15.3	57.7	18.3	91.6	7.0
	2010	58.1	15.9	58.1	17.9	92.2	6.1
	2011	55.4	17.3	57.1	19.4	86.1	9.0
	2012	54.7	17.1	57.2	19.4	86.3	9.4
	2013	58.9	17.2	56.9	19.3	89.3	7.8
	Decade Avg.	58.7	15.3	57.0	17.9	89.9	8.1
	RUNS (SPSS)	0.540		0.540			
Towns	2003	56.6	8.7	52.6	11.1	89.1	8.4
	2004	57.7	8.7	53.2	10.1	89.4	7.7
	2005	61.1	7.5	53.0	9.2	90.1	7.5
	2006	57.5	9.3	53.9	9.5	90.6	7.5
	2007	54.4	8.6	52.9	10.3	91.0	6.7
	2008	55.2	8.3	52.9	9.8	90.3	7.4
	2009	59.3	9.9	50.7	11.9	91.1	6.9
	2010	55.7	10.2	52.6	11.5	91.2	6.9
	2011	54.9	10.4	52.7	12.3	83.5	7.4
	2012	54.4	10.1	51.9	12.9	84.0	6.7
	2013	57.1	8.8	51.3	11.5	87.4	7.3
	Avg.	56.7	9.1	52.5	10.9	88.9	7.3
	RUNS (SPSS)	0.540		0.540			



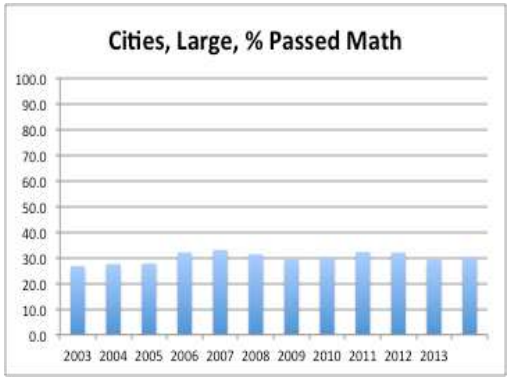
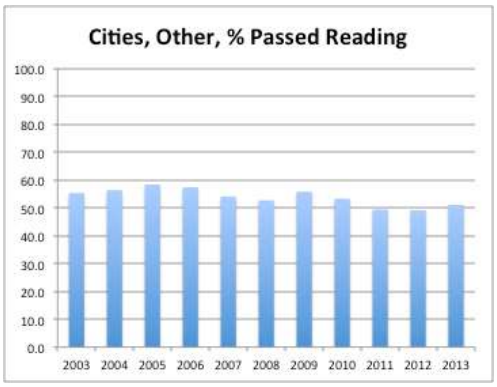


Figure 3: Reading and Math Average % Passed, Large Cities



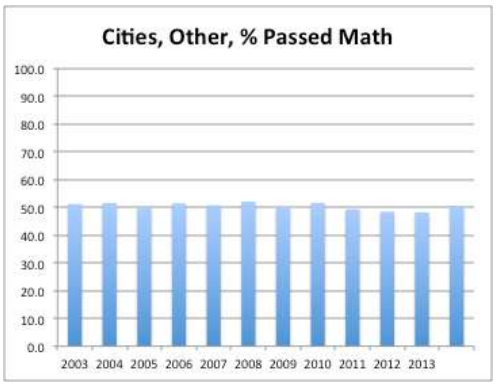
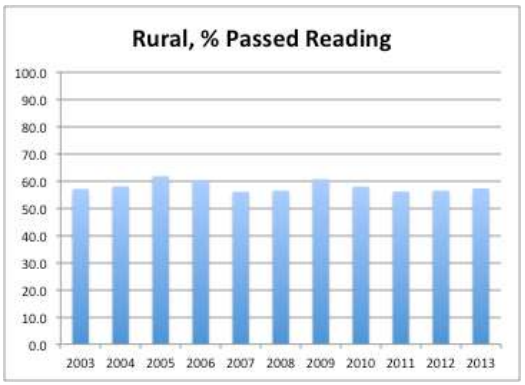


Figure 4: Reading and Math Average % Passed, Other Cities



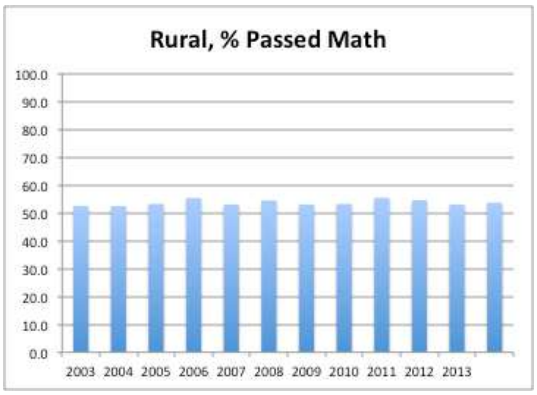
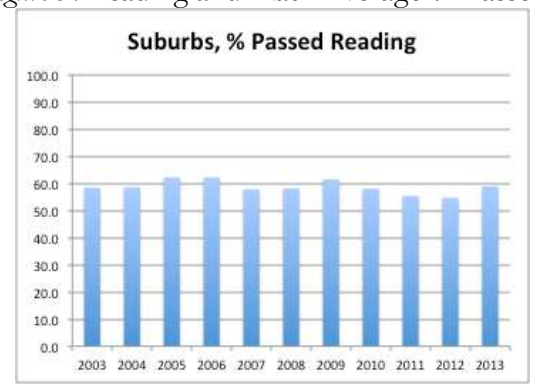


Figure 5: Reading and Math Average % Passed, Rural



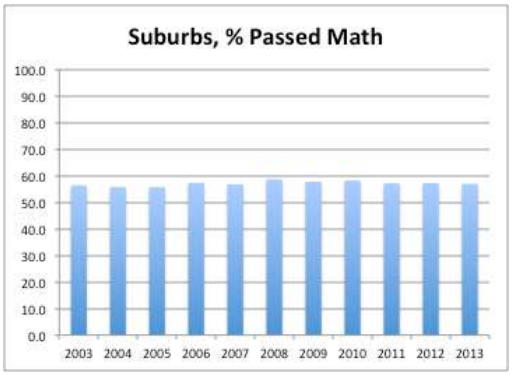
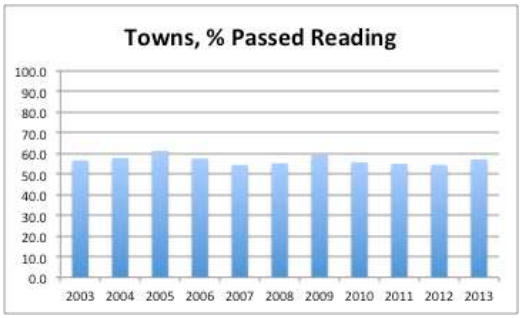


Figure 6: Reading and Math Average % Passed, Suburban



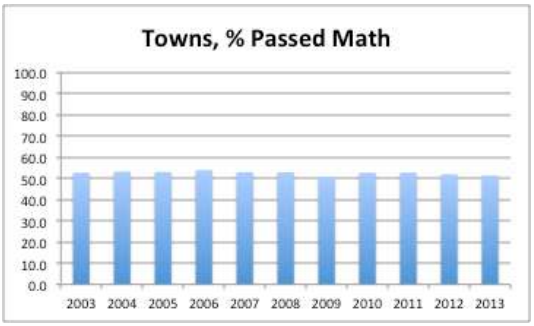


Figure 7: Reading and Math Average % Passed, Towns

Also, there was no significant variation in the graduation rates within schools statewide across the decade, with an average of 88.1% (standard deviation of 1.8). Nor was there any significant change in graduation rates by community type. (See Table 3 and Figure 8).

Table 3
Illinois Graduation Rates, 2003-2013, by State & by Community Type

	STATE	Cities Large	Other Cities	Rural	Suburbs	Towns
2003	88.4	72.3	86.9	91.7	90.2	89.1
2004	88.9	74.5	85.3	92.4	90.6	89.4
2005	89.1	76.3	85.8	92.1	90.4	90.1
2006	89.5	76.0	85.4	92.7	91.0	90.6
2007	88.7	68.9	86.3	93.0	90.4	91.0
2008	88.6	70.5	86.8	92.6	90.8	90.3
2009	89.0	71.3	86.8	93.1	91.6	91.1
2010	89.6	73.4	89.2	93.4	92.2	91.2
2011	84.4	77.4	81.9	87.3	86.1	83.5
2012	84.6	79.1	79.9	87.2	86.3	84.0
2013	88.6	86.8	81.0	91.2	89.3	87.4
Average	88.1	75.1	85.0	91.5	89.9	88.9
St.D.	1.8	5.0	2.9	2.2	2.0	2.8

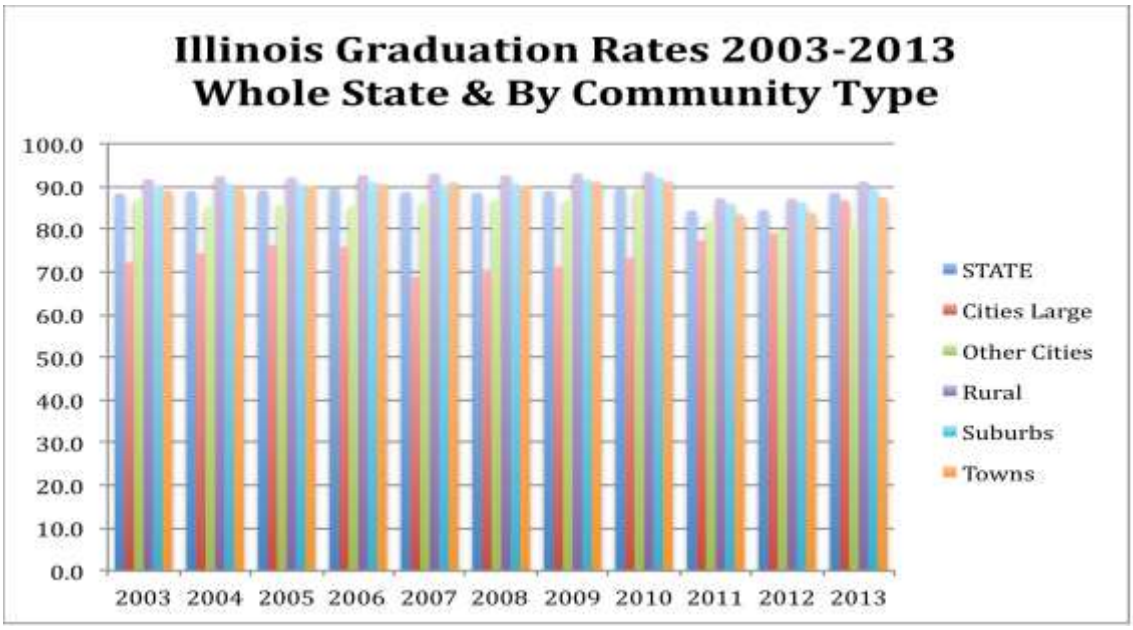


Figure 8: Illinois Graduation Rates 2003-2013, by State and by Community Type

Conclusions and Discussion

In Illinois, there was no significant gain in academic achievement in reading or mathematics as measured by the PSAE. Further, this result is not the result of gains by some communities versus declines experienced in others. Each community type – large city (Chicago plus the "suburb" of East Saint Louis), small city, town, rural, and suburban – remained essentially static.

This suggests that the program of focusing schooling on state mandated standards for core skill areas of reading and math, measured by standardized tests, is a failed experiment. The logic model of the standards movement as elaborated in the introduction seems to be inadequate at least at the second step: If teachers are held accountable for transmitting the high quality standards, then students will gain mastery of those standards. While accountability is likely to have been variable across the state, there was a lack of any significant progress over the decade in improving final test outcomes. Whether the incentives were negative, such as school closings and teacher firings, or were positive, such as receiving grants to aid progress, the effects were negligible.

It is possible that the failure was in the third step of the model that insists that standards align well with the assessments of the standards. It is possible that, in Illinois, the standards and the tests were not suitably aligned. This has been both suggested and refuted, as indicated in the data section. It seems unlikely that a lack of adequate alignment could explain such a static performance picture over such a span of years, especially given the proven tendency of teachers and schools to follow the performance elements and content on tests, rather than the standards themselves. It is also possible that the first, most philosophical and ideological component of the model is at the root of the problem, but that is far more speculative.

The breadth of the claim from this study, that the effects of the policy were negligible, could be argued on the basis that this research contemplates only the performance of high school juniors on their PSAEs. To counter this argument, note that the research on fourth and eighth grade outcomes indicates a similarly static picture. Moreover, measures in earlier grades were designed as benchmarks in terms of the final goal, which was meeting standards by high school graduation. If performance has not improved by the time they graduate high school, the overall mission has been a failure.

A major problem for standards/accountability based programs that emerged from the pursuit of the data for this study is logistical. The repetitions and holes in the data that the state included on the Illinois Report Card, seemingly without question, speak to the problems of administration and management of data which are vital for any high stakes approach. A detailed analysis of administrative capacity by Sunderman and Orfield (2006) described how the state agencies charged with overseeing, evaluating, and executing the provisions of NCLB were not provided with the expertise, infrastructure, number of employees, or technology sufficient to meet the new demands (pp. 7-13). Particularly in Illinois, they noted that, "the number of staff in the Illinois State Board of Education declined from 787 in FY 2000 to 492 in FY 2005, a 37.5% decrease during the time the state was implementing NCLB" (pp. 10-11). Inconsistencies in the dataset available for this study are likely a manifestation of the difficulties described by Sunderman and Orfield.

From the outset of NCLB, scholars in various states attempted to determine the additional spending that would be necessary for their states to meet AYP targets. The conclusion of all of them was that the amounts being put into place fell far short of the needed resources. An Education Policy Center study (EPC, 2005) pointed out that the

common underestimates of need coming from the GAO and state governments were a result of estimating compliance costs (the price of tests, for instance) rather than the much higher costs involved in increasing student performance (see Imazeki and Reschovsky, 2004; Mathis, 2005; Yinger, 2008). Illinois expenditures reflect this differentiation between compliance versus performance costs. Per pupil expenditure in Illinois rose 16% in real dollars between 2003-2013 (Illinois State Board of Education, 2013). This reflects attention to the scale of compliance cost increases, but not performance ones.

A related problem is that there has been no statewide mechanism to track district and school spending directed at performance improvement (i.e. teacher qualification improvements via hiring or professional development, additional staffing, new curricula and materials). The State only accounts for the increased compliance expenditures on tests, approximately $200 million between 2003-2013 (see Chingos, pp. 9, 25).

Another issue reflected in this research was noted by the Education Policy Study Center (2005): NCLB inevitably led to excluding many students because of where they live (pp. 6, 24-25). Rural schools that did not meet AYP for any of the years of the program seem from the data not to have received any of the consequences mandated by law. Unlike schools in Chicago that were held to the letter of the law to the point of closure and faculty/staff/school replacement, rural schools that failed remain continuous. This could be seen as having favored them or as having ignored their plight, and could be logistically unavoidable. Nonetheless, it represents unjust application of policy across the state.

It is also worth considering the unintended consequence of provisions legitimately designed to protect students' civil rights and privacy. By creating the cutoff of 40 students, approximately 18,500 students were not reported discretely in terms being members of minority populations in their schools. Thus, as Reardon et. al (2013, p. 5) point out, the students who we know might be struggling the most because of diversity issues are not accounted for, and their situation becomes as absent from the discussion as their performance is from the data.

There is also a lack of any growth in the connection between PSAE success and graduation rates. In every community, both PSAE passing rates and graduation rates remained static, but there is no apparent connection between the percentage of students passing their PSAEs and the percentage who were then allowed to graduate. In schools where PSAE passing rates could be as low as one-third of the students, the graduation rates were still approximately three-quarters of the students. This pattern was seen with some frequency in rural school districts in the state. While there was never any guarantee made by the laws that passing rates on basic skills would be connected to students' prospects for graduating, it does seem worth mentioning that many school districts graduated a lot of students who had not met the standards according to the PSAE metric.

From the limitations of any one study, it would be guessing to come to a firm conclusion about the reasons that there was no progress in passing rates. However, the researchers can suggest three possible reasons that have an impact upon future policy.

The first possibility is a failure in step three of the logic model, that the tests did not actually reflect performance on the desired skills well enough, or did not direct curricula to the skills as tested. This is the explanation currently favored by proponents of the Common Core and the accompanying PARCC test. According to this reason, the standards/accountability approach is the right direction, but we have not done it well enough. Improvements include greater precision in the metrics and in the actual skills being demanded by the standards. As suggested above, there is reason to suspect that this is not the real problem.

It should be noted that Illinois was faulted in various venues for selecting an ACT-based metric premised in evaluating college readiness. Yet the Common Core is targeted at college readiness. If Illinois was ahead of its time in terms of the goals, it suggests that the outcomes using the Common Core might end up reflecting Illinois' lack of success over the last decade. However, the situations are different, and the Common Core employs college readiness standards as well as metrics, which might be an improvement.

The second possibility stems from the first component of the logic model. It is possible that regular people and their communities do not actually desire for their children what is expected from the standards. The consistency of average scores across types of community suggests that this might be a reason for the failure of NCLB in Illinois. The idea that school should provide the same thing to every student, and that each student should emerge with the same skills as every other student, seems like a democratic approach. However, from within communities, the form and context for these skills may seem pointless to a life well lived. Schooling policy has proceeded from the assumption that the demands of university entrance, in the most reductionist terms of literacy and numeracy, should set the bar for what every student should know, and have never actually gathered any consensus of American communities on what is worthwhile.

Proponents of the Common Core claim that it was created by a consortium of states and included professional educators in its content and design. If irrelevance is a central issue in the failure of NCLB, it could be argued that such consortiums and professional educators are a tiny fraction of the stakeholders who need to be consulted to create effective curricula, standards, and assessments. No matter what schools demand, communities will be the ones to determine the actual outcomes.

The third possibility, also connected to the first step of the logic model, is that the entire paradigm could be incorrect. The assumptions about how learning happens and what knowing is could be in error. This research does not actually speak directly to this possibility, but from a future policy perspective, it would be remiss not to consider it as a possible reason for the failure of NCLB. The belief that the most valuable elements of an education are amenable to being serialized into atomized skills may be incorrect. What is known and widely accepted in neuropsychology and in education scholarship about human learning is that it is constructed, dialogic, and involves development of holistic meaning. Reading, critical thinking, creativity, and problem solving may be phenomenological rather than "skills based." If this is so, any skills approach to educating the population to be habitual and adept practitioners of such processes may be a dead end.

This study suggests that the expenditures in time, money, and human resources devoted to NCLB in the State of Illinois were largely wasted. The researchers recommend for future policy that tracking of expenditures aimed at meeting state program requirements be built in as part of any program initiated. The State would be wise to track all compliance and performance expenditures by districts in implementing the Common Core and accompanying assessments.

The other recommendations involve conversations. Forums need to be developed to inform policymakers about what is known about learning, pedagogy, and education rather than merely skill acquisition. Finally, conversations need to be structured and held such that every community gets voice and power in determining what their children get in school. Imposition of schooling policy from the top, employing assumptions about what comprises democracy and economic competitiveness for the nation, is unlikely to produce results any different from those seen in this report.

References

ACT and the Illinois State Board of Education. (2013). *Prairie State Achievement Examination Technical Manual.* Springfield, IL: Author. Retrieved from http://www.isbe.net/assessment/pdfs/psae/tech_manual13.pdf

Amrein-Beardsley, A. (2009). The unintended, pernicious consequences of "staying the course" in the United States' No Child Left Behind policy. *International Journal of Education Policy & Leadership, 4*(6), 1-12.

Blank, R. K. (2011). *Closing the Achievement Gap for Economically Disadvantaged Students? Analyzing Change Since No Child Left Behind Using State Assessments and the National Assessment of Educational Progress.* Washington, DC. Council of Chief State School Officers. Retrieved from http://eric.ed.gov/?id=ED518986

Bracey, G.W. (2005). No Child Left Behind: Where Does the Money Go? Tempe, AZ: Education Policy Studies Laboratory. Retrieved from http://files.eric.ed.gov/fulltext/ED508523.pdf

Chingos, M. M. (2012). *Strength in Numbers: State Spending on K-12 Assessment Systems.* Washington, DC: Brown Center on Education Policy at Brookings. Retrieved from http://www.brookings.edu/research/reports/2012/11/29-cost-of-ed-assessment-chingos

Duncombe, W., Lukemeyer, A., & Yinger, J. (2008). The No Child Left Behind Act: Have federal funds been left behind? *Public Finance Review, 36*(4), 381- 407. http://dx.doi.org/10.1177/1091142107305220

Haladyna, T.M., Nolen, S.B., & Haas, N.S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(5), pp. 2-7. http://dx.doi.org/10.3102/0013189X020005002

Illinois State Board of Education (2012). *A Profile of Illinoi Public Schools: Selections from the School Report Card Files.* Retrieved from http://www.isbe.net/reports/annual12/toc.htm

Koretz, D.N. (2005). *Alignment, high stakes, and the inflation of test scores. CSE Report 655.* Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE), Graduate School of Education & Information Studies.

Lawrence, C.R. III (2006). Who Is the Child Left Behind? The Racial Meaning of the New School Reform. Georgetown University Law Center, Scholarship@ Georgetown Law. Retrieved from http://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article= 1338&context=facpub

Lee, J & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990–2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, *34*(2), pp. 209-231. http://dx.doi.org/10.3102/0162373711431604

Mathis, W.J. (2005). The cost of implementing the federal No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education, 80*(2), pp. 90-119. http://dx.doi.org/10.1207/S15327930pje8002_6

National Center for Education Statistics (2013). The Condition of Education 2013. Washington, DC: Author Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2013037

National Council on Disability (2008). *The No Child Left Behind Act and the Individuals with Disabilities Education Act: A Progress Report.* Washington, DC: Author.

PSAE Trend Data 2006-214. Retrieved from http://www.isbe.net/assessment/report_card.htm

Ponisciak, S.M. (2005). *Understanding the Prairie State Achievement Exam*. Chicago, IL: Consortium on Chicago School Research, University of Chicago. Retrieved from https://ccsr.uchicago.edu/publications/understanding-prairie-state-achievement-exam-descriptive-report-analysis-student

Reardon, S.F., Greenberg, E., Kalogrides, D., Shores, K.A., & Valentino, R.A. (2013). *Left Behind? The Effect of No Child Left Behind on Academic Achievement Gaps*. Stanford, CA: Center for Educational Policy Analysis. Retrieved from http://cepa.stanford.edu/content/left-behind-effect-no-child-left-behind-academic-achievement-gaps

Sims, D.P. (2007). Can failure succeed? Using racial subgroup rules to analyze vouchers, stigma, and school accountability. Unpublished working paper. Brigham Young University, Department of Economics.

Srikantaiah, D. (2009). How State and Federal Accountability Policies Have Influenced Curriculum and Instruction in Three States. Washington, DC: Center on Education Policy. Retrieved from http://www.cepdc.org/displayDocument.cfm?DocumentID=217

Sunderman, G. L. & Orfield, G. (2006). Massive responsibilities and limited resources: The state response to NCLB. Paper presented at the NCLB Key Issues: Reforming AYP and Evaluating State Capacity roundtable co-sponsored by The Civil Rights Project at Harvard University and The Chief Justice Earl Warren Institute on Race, Ethnicity and Diversity at the University of California, Berkeley, School of Law, Washington, DC, November 16-17. Retrieved from http://www.academia.edu/3146090/Massive_Responsibilities_and_Limited_Resources_The_State_Response_to_NCLB1

About the Authors

Wm. Gregory Harman
Dominican University, Chicago
wharman@dom.edu

Greg Harman, Ed.D., is an Assistant Professor in his sixth year in the School of Education at Dominican University. He coordinates the secondary, social studies, and Teach for America secondary programs. He has published in *Teaching & Learning: The Journal of Natural Inquiry & Reflective Practice*, *Teaching with Primary Sources*, and *Critical Questions in Education*. He has presented at the University of Illinois Forum on the Future of Public Education and at the Critical Questions in Education Conference.

Camille Boden
Dominican University, Chicago
bodecami@my.dom.edu

Camille Boden, MAT, is a kindergarten teacher at Rowe Elementary School in Chicago. She has experience as an English teacher in Lyon, France, and as a reading teacher for the Institute of Reading Development. She presented at the Critical Questions in Education Conference in Louisville, Kentucky in the fall of 2014.

Jeremy Karpenski
Dominican University, Chicago
karpjere@my.dom.edu
Jeremy Karpenski, MAT, is a middle school social studies teacher at East Prairie School in Skokie, IL. He presented at the Critical Questions in Education Conference in Louisville, Kentucky in the fall of 2014.

Nicole Muchowicz
Dominican University, Chicago
muchnico@my.dom.edu
Nicole Muchowicz is currently in the Master of Arts in Teaching program at Dominican University. She is also a full time English teacher and the Theatre Director at Saint Joseph High School in Westchester, Illinois.

education policy analysis archives

Volume 24 Number 48 April 18, 2016 ISSN 1068-2341

archivos analíticos de políticas educativas
consejo editorial