

State of New Mexico
LEGISLATIVE EDUCATION STUDY COMMITTEE

REPRESENTATIVES

Rick Miera, Vice Chair
Joni Marie Gutierrez
Dianne Miller Hamilton
Mimi Stewart
Thomas E. Swisstack
W. C. "Dub" Williams

State Capitol North, 325 Don Gaspar, Suite 200
Santa Fe, New Mexico 87501
PH: (505) 986-4591 FAX: (505) 986-4338
<http://legis.state.nm.us/lcs/lesc/lescdefault.asp>

SENATORS

Cynthia Nava, Chair
Gay G. Kernan
Mary Kay Papen
William E. Sharer

ADVISORY

Ray Begaye
William "Ed" Boykin
Kandy Cordova
Roberto J. Gonzales
Jimmie C. Hall
John A. Heaton
Harriet I. Ruiz
Sheryl M. Williams Stapleton
Richard D. Vigil
Teresa A. Zanetti



ADVISORY

Vernon D. Asbill
Mark Boitano
Carlos R. Cisneros
Dianna J. Duran
Mary Jane M. Garcia
John Pinto
Leonard Tsosie

D. Pauline Rindone, Ph.D., Director
Frances R. Maestas, Deputy Director

October 17, 2006

MEMORANDUM

TO: Legislative Education Study Committee

FR: Pamela Herman

**RE: STAFF BRIEF: ASSESSMENT AND ACCOUNTABILITY: STATE
ASSESSMENTS AND NAEP: DISPARITY IN TEST RESULTS**

The 2006 interim workplan of the Legislative Education Study Committee (LESC) includes a presentation examining the disparity in test results between the New Mexico Standards Based Assessments administered under the *Assessment and Accountability Act* and the National Assessment of Educational Progress administered to students in every state as required by federal law.

Issues:

The federal *No Child Left Behind Act of 2001* (NCLB) requires every state that accepts Title I funds to develop and follow a plan that includes annual testing in grades 3 through 8 and at least once in high school based on state standards in reading/language arts and mathematics; by school year 2007-2008, students must also be tested once in elementary school, in middle school, and in high school in science. States must also assess the English proficiency of English-language learners (ELLs) annually and administer the National Assessment of Educational Progress (NAEP) to a sample of students in grades 4 and 8 in alternate years.

In 2003, the LESL endorsed and the Legislature passed the *Assessment and Accountability Act* as a component of a comprehensive package of school reforms designed to conform to NCLB. State statute requires the Public Education Department (PED) to establish a statewide assessment and accountability system aligned with state academic content and performance standards. PED was required to begin testing student achievement for grades 3 through 9 and grade 11 in mathematics, reading and language arts, and social studies if funds were available by school year

2005-2006; and for grades 3 through 5, 6 through 9, and 11 in science by school year 2007-2008; and was required to apply writing assessment scoring criteria to the extended response writing portions of language arts standards-based assessments in grades 3 through 9 by school year 2005-2006.

PED has implemented a system of standards-based assessments that complies with, and where state requirements dictate, in some cases exceeds, the requirements of NCLB. The New Mexico State Accountability Plan to provide these assessments is fully approved by the US Department of Education (USDE); however, the state standards and assessments themselves are subject to a federal peer review process, and as of June 2006 the status of New Mexico's standards and assessment system was deemed "Approval Pending" by USDE. New Mexico is under Mandatory Oversight while it complies with a timeline to provide USDE with additional evidence supporting certain aspects of the assessment system by the end of the current school year.¹

School year 2004-2005 was the first for which results were available for both the New Mexico standards-based assessments and the NAEP (see Attachment 1). A comparison of results of the state and federal assessments shows the following:

- The percentage of students of various groups who scored proficient or better on the NAEP was lower than on the state assessments by the following margins:
 - 20 to 34 points for grade 4 in reading;
 - 27 to 38 points in grade 8 reading;
 - 10 to 24 points in grade 4 mathematics; and
 - 4 to 14 points in grade 8 mathematics.
- The achievement gaps based on race and ethnicity, income, English language learner status, and gender were not always comparable on both assessments.

According to School Matters, a service of Standard & Poors (S&P), the increased attention on standardized tests focused by NCLB has resulted in public confusion about the relationship between the NAEP and state assessments, especially where a state's performance varies significantly on the different exams. S&P suggests that while such differences may lead to valid conclusions that state assessments are insufficiently rigorous, there are key differences between the exams that may contribute to a performance gap on the two systems. These differences include the following:

- There are no consequences attached to student (or, for that matter, school) performance on the NAEP;
- No instructional hours are spent specifically preparing for the NAEP;
- The NAEP is administered only to a "stratified random sample" of students – according to NCES, 2,500 students in approximately 100 schools per grade and per subject in an average state;
- The NAEP and state assessments have different provisions for inclusion of students with cognitive or language differences, and NAEP guidelines for accommodations prohibit translation of NAEP reading exams into other languages;

¹ For a more comprehensive description of state and federal assessment requirements, see the October 2006 LESC staff brief "Assessment and Accountability: NCLB and State Assessment Requirements."

- NAEP assessments are based on content frameworks and specifications developed by a national board, while state assessments are aligned to state academic content standards that are not required to be, and in fact may not be, aligned with the federal framework;
- The range of performance levels to be considered proficient (or less than proficient, or advanced) may vary on the NAEP and state exams; for the NAEP, proficiency indicates “demonstrated competency over challenging subject matter”; and
- The number of questions a student must answer correctly (the “cut score”) may differ for the NAEP and the state exams. (See Attachment 2 for a summary of NAEP proficiency levels and 4th grade reading cut scores.)

S&P states that a state’s performance gap on the two exams should not automatically be interpreted as an indicator of the relative rigor of either test. If a gap exists between a state’s performance on the state assessments and the NAEP, policymakers should determine the degree to which the factors listed above influence their state’s performance, and the manner in which these issues should best be addressed by the state. S&P concludes that evaluating both tests remains a useful exercise for policymakers who wish to understand trends in their state’s overall academic performance and the academic progress of specific subgroups of students within the state.

Background:

The NAEP

The NAEP, also called The Nation’s Report Card, is a congressionally mandated project of the National Center for Education Statistics (NCES) within the USDE. According to NCES, since its inception in 1969 the NAEP has measured the nation’s educational progress by regularly administering various subject-area assessments to nationally representative samples of students. In federal statute, the NAEP collectively refers to a national assessment, state assessments, and a long-term trend assessment in reading and mathematics.

In 1988, Congress established the National Assessment Governing Board (NAGB) to oversee and set policy for the NAEP. The board is responsible for the following:

- selecting the subject areas to be assessed;
- setting appropriate student achievement levels;
- developing assessment objectives and test specifications;
- developing a process for the review of the assessment;
- designing the assessment methodology;
- developing guidelines for reporting and disseminating NAEP results;
- developing standards and procedures for interstate, regional and national comparisons;
- determining the appropriateness of all assessment items and ensuring the assessment items are free from bias and are secular, neutral and non-ideological;
- taking actions to improve the form, content, use and reporting of results on the NAEP; and planning and executing the initial public release of NAEP reports.

Until the passage of NCLB, participation in the NAEP was voluntary for students, schools, local school districts and states. In 2001, however, Congress mandated that states receiving Title I funds must participate in biennial reading and mathematics assessments in grades 4 and 8.

NCLB requires the US Secretary of Education to conduct a national assessment of the programs funded under Title I and the impact of Title I on states, local school districts, schools and students, using information from a variety of sources “including the National Assessment of Educational Progress. . . State evaluations, and other research studies.” The Secretary is required to transmit an interim report of this national assessment to the President and Congress not later than three years after the enactment of NCLB, or January 2004, and a final report not later than five years after enactment, or January 2006. However, the interim report was not published until February 2006. Regarding national trends in student achievement data since the enactment of NCLB, the interim report states the following:

- In states that had three-year trend data available from school years 2000-2001 to 2002-2003, the percentage of students achieving at or above the state’s proficient level rose for most student subgroups in the majority of states but the increases in student proficiency were often small.
- Recent NAEP trends show gains in 4th-grade reading and especially in mathematics for black and Hispanic students and for students in high poverty schools.
- State assessments and NAEP both provide some indications that achievement gaps between disadvantaged students and other students may be narrowing, but recent changes are small.

State Standards Based Assessments

The Education Commission of the States (ECS) attributes the initiation of the modern standards movement in education to the National Council on Education Standards and Testing, established by Congress in 1991, which issued a report the following year calling for the development of national standards in each of the major subject areas, embodying “demanding but attainable learning goals” for the widest possible range of students. While the debate about national standards continues, by the time Congress passed NCLB many states including New Mexico had established their own standards in most subjects.

In 2000, the New Mexico Education Initiatives and Accountability Task Force (EIATF), convened by the Speaker of the House of Representatives and the President of the Senate to review the status of the New Mexico education system and study possible reforms, made its final report to the Legislature. Among its recommendations, the task force stated that “It is a state-level responsibility to set statewide performance-based standards for student achievement, which will be measured annually; districts and schools will be held responsible for the results of those annual measurements.” In 2003, the LESC endorsed and the Legislature passed HB 212, *Public School Reforms*, to embody in statute many of the reforms recommended by the EIATF and the LESC Ad Hoc Subcommittee for Education Reform regarding inclusion of a statewide educational accountability system built on uniform state standards-based assessments.

Presenters:

Dr. Andrew Kolstad, Senior Technical Advisor, Assessment Division, NCES, will provide the committee with an overview of the NAEP assessment system for mathematics and reading in 4th and 8th grade, including performance trends in New Mexico and nationally.

Dr. Don Watson, Assistant Secretary, Assessment and Accountability Division, PED, will describe the process PED used to develop state standards-based assessments and establish proficiency levels. Dr. Karen K. Harvey, Assistant Secretary, Quality Assurance and Systems Integration, PED, will discuss the issue of the alignment of state assessments with the NAEP.

Questions the committee may wish to consider:

1. What is the definition of “proficiency” on standards-based assessments in New Mexico?
2. To what extent is the sample of students who take the NAEP in New Mexico representative of the composition of New Mexico’s public school population?
3. What, if anything, does a comparison of the performance of New Mexico students on the NAEP and the NM standards-based assessments indicate about student academic achievement?
4. What further analysis, if any, might be warranted to evaluate what the differences in student performance on the exams indicates about the exams themselves?
5. To what extent might it be appropriate for New Mexico to have academic standards that differ in significant ways from the NAEP assessment framework, or cut scores that measure proficiency differently?
6. What relationship, if any, exists between the NAEP assessment framework and the standards embodied in other nationally administered exam including college admissions examinations such as ACT and SAT?

Comparison of NM Standards Based Assessments (SBA)
and NAEP Results: Spring 2005

READING									
		New Mexico SBA Results ¹				New Mexico NAEP Results			
		% Begin. Prof.	% Near Prof.	% Prof.	% Adv.	% Below Basic	% Basic	% Prof.	% Adv.
Grade 4	All Students	11	35	42	10	49	31	17	4
	Asian	3	23	49	23	Reporting standards not met			
	Black	14	39	37	8	50	26	21	4
	Hispanic	13	40	38	7	57	29	12	2
	Native Amer	19	46	29	3	67	24	8	<1
	White	5	23	53	17	28	36	28	8
	Male	13	37	40	8	53	30	14	3
	Female	9	34	44	12	44	32	19	5
	Low Income	14	41	37	6	58	29	12	1
	Eng Learner	18	47	29	4	76	19	5	<1
Grade 8	All Students	11	36	48	3	38	43	18	1
	Asian	5	21	59	11	Reporting standards not met			
	Black	13	40	44	2	Reporting standards not met			
	Hispanic	13	40	44	2	45	43	12	1
	Native Amer	16	47	34	1	51	42	6	1
	White	6	26	61	6	24	43	31	2
	Male	15	40	42	2	43	40	16	1
	Female	7	33	55	5	33	45	21	1
	Low Income	14	42	41	2	46	42	12	<1
	Eng Learner	18	48	31	1	70	27	3	1
MATHEMATICS									
Grade 4	All Students	10	50	32	7	35	46	17	2
	Asian	5	25	45	24	Reporting standards not met			
	Black	13	56	26	4	55	39	6	<1
	Hispanic	12	55	28	5	43	44	12	1
	Native Amer	14	59	22	3	44	48	8	<1
	White	5	39	42	13	17	48	30	4
	Male	10	49	32	8	35	44	19	2
	Female	9	51	32	7	36	47	16	1
	Low Income	12	56	27	4	43	45	11	<1
	Eng Learner	15	60	22	3	58	37	5	<1
Grade 8	All Students	24	51	20	4	47	39	13	1
	Asian	11	34	36	18	Reporting standards not met			
	Black	31	51	15	2	56	31	13	<1
	Hispanic	29	53	15	2	57	35	7	1
	Native Amer	34	53	10	1	61	35	4	<1
	White	13	46	32	8	28	46	23	3
	Male	26	48	20	4	47	39	13	2
	Female	22	53	20	4	48	39	12	1
	Low Income	31	53	13	1	59	34	7	<1
	Eng Learner	38	51	8	1	77	21	2	<1

Source: PED and NCES

¹ NM SBA totals do not equal 100% because of invalid tests

SchoolMatters

A Service of STANDARD & POOR'S

Appendix B

NAEP

NAEP Proficiency Levels:

- **Advanced:** Superior Performance.
- **Proficient:** Solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situation, and analytical skills appropriate to the subject matter.
- **Basic:** Partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

2003 NAEP "Cut Scores"		
Proficiency Level	4 th Grade Reading	4 th Grade Math
Advanced	268	282
Proficient	238	249
Basic	208	214

NAEP 2003 Grade 4 Reading Scale

The chart below illustrates the knowledge and skills demonstrated by students performing at different scale scores on the 2003 NAEP reading assessment. The scale score associated with each question represents the probability that, at any given score point, 65% of the students (for a constructed-response question) and 74% of the students (for a four-option multiple-choice question) answered that question successfully. For constructed-response questions, responses could be completely or partially correct and therefore a question can map to several points on the scale.

For example, in the case of the multiple-choice question that maps at 172 on the scale, 4th grade students with a score of 172 have a 74% chance of answering this question correctly. In other words, out of a sample of 100 students who scored 172, 74 would be expected to have answered this question correctly.

2003 NAEP Reading Scale

*MC = Multiple Choice, CR = Constructed Response

500	
360	
352	Extend relevant information to make an inference. (CR)
350	
340	
330	
322	Explain causal relation between pieces of text information. (CR)
320	
319	Use metaphor to compare story characters. (CR)
310	
301	Describe character's changing feelings and explain cause. (CR)
300	
294	Provide and explain an alternative ending to a story. (CR)
290	
286	Provide alternative title and support with story details. (CR)
280	
270	Explain the author's use of direct quotations. (CR)
269	Use character trait to compare to prior knowledge. (CR)
268	Advanced
266	Provide overall message of story. (CR)
262	Explain author's statement with text information. (CR)

SchoolMatters

A Service of STANDARD & POOR'S

- 260
257 Discriminate between closely related ideas. (MC)
255 Make inference to identify character motivation. (MC)
250 Retrieve relevant information to fit description. (CR)
245 Provide a cause for character's emotion. (CR)
240 Identify explicit embedded information related to main topic. (MC)
Provide text-based lesson. (CR)
239 Identify main theme of story. (MC)
238 **Proficient**
232 Retrieve text details to make a comparison. (CR)
230 Use prior knowledge to make text-related comparison. (CR)
226 Recognize main reason that supports idea. (MC)
221 Recognize meaning of specialized vocabulary from context. (MC)
220
214 Retrieve text details to provide a description. (CR)
213 Provide text-based inference. (CR)
210 Recognize text-based inference. (MC)
208 **Basic**
200
196 Retrieve and provide a text-related fact. (CR)
190
180
179 Recognize story type as adventure. (MC)
172 Identify character's main dilemma. (MC)
170
160
150
145 Recognize explicit fact repeated across text. (MC)
0

Sources:

<http://nces.ed.gov/nationsreportcard/itemmaps/index.asp?grade=4&subj=Reading>
<http://nces.ed.gov/nationsreportcard/itemmaps/index.asp>
<http://nces.ed.gov/nationsreportcard/reading/achieve.asp>