



**Report
to
The LEGISLATIVE FINANCE COMMITTEE**



**Public Education Department
Program Evaluation of State Assessment Program
January 15, 2009**

Report #09-01

LEGISLATIVE FINANCE COMMITTEE

Senator John Arthur Smith, Chairman
Representative Luciano “Lucky” Varela
Senator Sue Wilson Beffort
Representative Donald E. Bratton
Senator Pete Campos
Senator Carlos R. Cisneros
Senator Phil A. Griego
Senator Stuart Ingle
Representative Rhonda S. King
Representative Brian K. Moore
Senator Mary Kay Papen
Senator Leonard Lee Rawson
Representative Henry “Kiki” Saavedra
Representative Nick L. Salazar
Representative Edward C. Sandoval
Representative Jeannette O. Wallace

DIRECTOR

David Abbey

DEPUTY DIRECTOR FOR PROGRAM EVALUATION

Manu Patel, CPA

PROGRAM EVALUATION TEAM

David Craig
Lawrence Davis
Brenda Fresquez, CICA
Pam Galbraith
Donna Hill-Todd, CGFM
Craig Johnson
John Ketchens
Jordan Maril
Charles Sallee
Aurora B. Sánchez, CISA

SENATOR JOHN ARTHUR SMITH
CHAIRMAN

Senator Sue Wilson Beffort
Senator Pete Campos
Senator Carlos R. Cisneros
Senator Phil A. Griego
Senator Stuart Ingle
Senator Mary Kay Papen
Senator Leonard Lee Rawson

State of New Mexico
LEGISLATIVE FINANCE COMMITTEE

325 Don Gaspar, Suite 101 • Santa Fe, New Mexico 87501
Phone: (505) 986-4550 • Fax: (505) 986-4545

DAVID ABBEY
DIRECTOR



REPRESENTATIVE LUCIANO "LUCKY" VARELA
VICE-CHAIRMAN

Representative Donald E. Bratton
Representative Rhonda S. King
Representative Brian K. Moore
Representative Henry "Kiki" Saavedra
Representative Nick L. Salazar
Representative Edward C. Sandoval
Representative Jeannette O. Wallace

January 15, 2009

Dr. Veronica Garcia, Secretary
New Mexico Public Education Department
Jerry Apodaca Education Building
300 Don Gaspar
Santa Fe, NM 87501

Dear Dr. Garcia,

On behalf of the Legislative Finance Committee (Committee), I am please to transmit the *Program Evaluation of the State Assessment Program*.

The review team evaluated the performance and oversight of the assessment program including a cost benefit analysis of the current standards based assessment and an examination of the Public Education Department's contract management practices. The report will be presented to the Committee on January 15, 2009. An exit conference was conducted on January 6, 2009 and on January 12, 2009 to discuss the contents of the report with you and your respective staff.

I believe this report addresses issues the Committee asked us to review and hope your department benefits from our efforts. We very much appreciate the cooperation and assistance we received from you and your staff.

Sincerely,

A handwritten signature in blue ink that reads "David Abbey".

David Abbey, Director

DA:CJ/svb

Table of Contents

	<u>Page No.</u>
EXECUTIVE SUMMARY	1
BACKGROUND INFORMATION	3
FINDINGS AND RECOMMENDATIONS	5
NEW MEXICO HAS A RIGOROUS, BUT EXPENSIVE STANDARDS BASED ASSESSMENT. OPPORTUNITIES EXIST TO LOWER COSTS.....	5
APPROPRIATIONS TO PED FOR ASSESSMENT HAVE BEEN SUBSTANTIAL.....	10
PED’s INADEQUATE CONTRACT MANAGEMENT RESULTS IN EXCESSIVE COSTS	13
AGENCY RESPONSES.....	16
APPENDIX A: NEW MEXICO STATE ASSESSMENT PROGRAM	24

The current contract for the New Mexico Standards Based Assessment expires in June 2009. Total NMSBA contract costs to PED and districts are about \$31.1 million.

Constructed response questions increase assessment costs and take longer to score.

Other states have developed rigorous assessments that use constructed response less.

Appropriations for public education in New Mexico amount to about \$2.6 billion annually and represent about 43 percent of state spending. General fund appropriations to the New Mexico Public Education Department (PED) for FY09 operations were \$17.1 million, a 14.6 percent increase over FY08. Over the past few years, the state has implemented several education reforms to improve student achievement. PED's assessment program manages the key assessments designed to measure progress in student achievement.

The Legislature's significant investments in public education include substantial funding for assessments. The purpose of the Assessment and Accountability Act (Chapter 22 Article 2C NMSA 1978) is to comply with accountability requirements established under the federal No Child Left Behind Act (NCLB). Statutory duties outlined in the Assessment and Accountability Act require PED to establish a statewide assessment and accountability system that is aligned with state performance standards.

PED currently contracts with NCS Pearson to develop and score the New Mexico Standards Based Assessment (NMSBA). Total costs to PED and districts from FY06 to FY09 under this contract, which expires on June 30, 2009, are about \$31.1 million. PED anticipates releasing a Request for Proposal (RFP) for the next NMSBA contract in early 2009. This review evaluates the performance and oversight of the assessment program including a cost-benefit analysis of the current standards based assessment and an examination of PED's contract management practices.

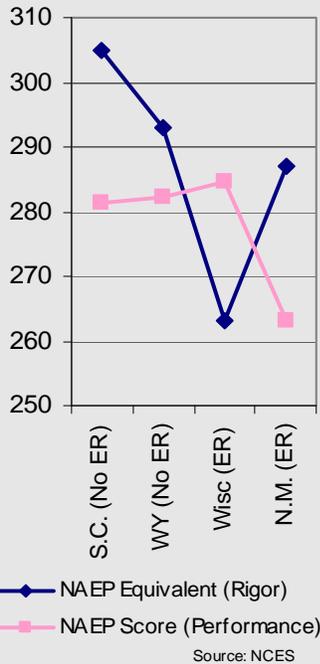
Student achievement is a top state priority and limited resources could be reallocated from activities that measure student achievement to those that improve student achievement.

Key Findings

New Mexico has a rigorous but expensive standards based assessment. Opportunities exist to lower costs. Use of constructed response questions significantly increases assessment costs. Multiple choice questions can be graded by machine, but constructed response questions are scored by hand. Constructed response questions increase the time needed to score assessments. This delays results and can create reliability issues. Opportunities exist to lower the cost of state assessments without an unacceptable loss of rigor. Other states have developed assessment systems that rely less on constructed response and still maintained rigor.

The National Center for Education Statistics (NCES) did not find a strong relationship between state assessment rigor and performance on national tests.

8th grade Math-2005



PED's ability to manage contracts needs improvement.

Appropriations to PED for assessment have been substantial. From Federal Fiscal Year (FFY) 05 to FFY08, federal appropriations to PED for the development and implementation of assessments were over \$18 million. From FY06 through FY09, the Legislature appropriated \$19.6 million to PED for development and implementation of state assessments. Over roughly the same period of time, PED expenses for the three main assessments were about \$28.5 million or about 75 percent of the amount of state and federal funding. The Legislature appropriated more than \$7.8 million in additional recurring dollars through the Public School Funding Formula to provide for school district assessment costs. PED provided information on January 9, 2009 describing assessment related expenditures; however, this information has not yet been fully verified by LFC staff.

PED's inadequate contract management results in excessive costs. PED contract management is inadequate. PED is not effectively tracking deliverables or contractor performance. Contracts contain vague language regarding PED responsibilities and there have been several procurement code violations.

Key Recommendations

Save money by reducing the use of constructed response in certain areas. PED should write the RFP for the next NMSBA contract to direct bidders to submit proposals with various levels of constructed response and expected time frames for delivery of results. Develop assessment contracts to make test results from constructed response questions available before the next school year and results from multiple choice questions available in the year the test is administered. This would enable districts and schools to make end of year adjustments.

PED should provide additional documentation on how all appropriations for assessment have been spent. The federal appropriation level should serve as a target for NMSBA contract expenditures.

PED must improve their ability to effectively administer contracts. Improved contract monitoring and tracking of deliverables would reduce procurement code violations. Future deliverables should include clear measures of quality.

Share the expense of developing high-cost, low-use assessments with neighboring states.

BACKGROUND INFORMATION

With the federal No Child Left Behind Act of 2001 (NCLB), Congress increased the frequency with which states measure student achievement and authorized funding to support state efforts to develop and implement these tests. NCLB requires annual assessments in math and reading/language arts aligned to content standards and administered to all students in grades 3-8 and a grade level in high school. Assessments in science are required in at least one grade in elementary school (grades 3 through 5), one grade in middle school (grades 6 through 9) and one grade in high school (grades 10 through 12). This amounts to at least 17 NCLB tests each year for school districts and entails a considerable amount of school time for administering assessments.

The New Mexico Public Education Department (PED) administers the state's assessment program which establishes proficiency levels and is approved by the U.S. Secretary of Education. This evaluation describes characteristics of New Mexico's Standards Based Assessment (NMSBA), identifies factors that explain variation in expenses, illustrates funding for assessment, and reviews PED's management of contracts. It is anticipated that approximately 169,000 students in grades 3 through 8 and 11 will take the NMSBA in spring 2009. See appendix for a description of New Mexico's key assessments.

<u>Year</u>	<u>Number of test takers</u>	<u>Grades</u>
2007-2008	166,085	grades 3-8, 11
2006-2007	193,995	grades 3-9, 11
2005-2006	196,024	grades 3-9, 11
2004-2005	196,458	grades 3-9, 11

TESTING COMPANIES

PED has contracted with NCS Pearson (Pearson) to develop and score the NMSBA, the New Mexico High School Competency Exam (NMHSCE), and New Mexico English Language Proficiency Assessment (NMELPA). About a year ago, Pearson completed its acquisition of Harcourt Assessment. American Institutes of Research (AIR) has the contract for the New Mexico Alternate Performance Assessment (NMAPA).

NCLB created business opportunity for testing companies. Some states require tests in addition to NCLB requirements. Several companies have the capacity to develop, administer, and score the NMSBA. Other prominent testing companies include CTB-McGraw Hill, Riverside Publishing (Houghton Mifflin), Data Recognition, Educational Testing Service (ETS), Measured Progress, Measurement Incorporated, WestEd, Westat, and Questar Assessment.

TYPES OF ASSESSMENTS

Types of assessments include formative and summative assessments. Formative assessments can be used to target individual students and modify day to day classroom activities whereas summative assessments, like the NMSBA, are better suited to identify state-wide, district-wide, or school-wide weaknesses and alter instructional focus. Types of questions on assessments can be multiple choice (MC) or constructed response, which include fill in the blank (FTB), short answer (SA), and extended response (ER) questions. Extended response questions require students to write an essay or demonstrate how they arrived at an answer.

PROJECT INFORMATION

Program Evaluation Objective. The objective of the evaluation is to review the performance and oversight of the standards based assessment program including a cost-benefit analysis of the current assessments and a review of PED's contract management.

Program Evaluation Activities (Scope and Methodology).

- Review statutes to develop an understanding of assessment requirements,
- Identify potential cost savings and best practices associated with testing contracts,
- Benchmark New Mexico's Standards Based Assessment with other states assessments in terms of item response types, timeliness of results, and related costs,
- Review appropriations and expenditures for assessment,
- Review documentation reflective of PED's administration of contracts.

Authority for Review. The LFC has the statutory authority under Section 2-5-3 NMSA 1978 to examine laws governing the finances and operations of departments, agencies and institutions of New Mexico and all of its political subdivisions, the effects of laws on the proper functioning of these governmental units and the policies and costs. The LFC is also authorized to make recommendations for change to the Legislature. In furtherance of its statutory responsibility, the LFC may conduct inquiries into specific transactions affecting the operating policies and cost of governmental units and their compliance with state law.

Program Evaluation Team.

Craig Johnson, Project Lead

David Craig, Program Evaluator

Exit Conference. The contents of this report were discussed with senior staff from the Public Education Department on January 6, 2009 and again with Secretary Garcia and PED staff on January 12th, 2009.

Report Distribution. This report is intended for the information of the Office of the Governor, Department of Finance and Administration, Public Education Department, the Office of the State Auditor, and the Legislative Finance Committee. This restriction is not intended to limit distribution of this report which is a matter of public record.



Manu Patel

Deputy Director for Program Evaluation

FINDINGS AND RECOMMENDATIONS

NEW MEXICO HAS A RIGOROUS BUT EXPENSIVE STANDARDS BASED ASSESSMENT. OPPORTUNITIES EXIST TO LOWER COSTS.

New Mexico uses one of the more expensive assessments. Assessment costs include development, administration, scoring and reporting. Extensive use of constructed response questions and developing a Spanish version of the NMSBA add to the costs. The bulk of costs associated with extended response are related to scoring. Future development costs should be reduced as PED owns the questions already developed and the NMSBA will serve as a graduation exam beginning in FY11.

Variation in assessments and contracts makes cost comparisons complicated. New Mexico has one of the highest costs per assessment according to an LFC survey and a GAO report. The cost data in Table 1 reflect that use of constructed response significantly increases expense; however

Table 1. Estimated Assessment Costs

State	Estimated cost per assessment	Multiple choice only
New Mexico	\$18.14	N
Utah	\$3.52	Y
Idaho	\$6.00	Y
North Carolina*	\$1.85	Y
Texas*	\$4.72	Y
Virginia*	\$1.80	Y
Nevada	\$12.00	N
Minnesota	\$14.32	N
Florida	\$15.14	N
Vermont	\$15.50	N
Washington	\$20.00	N
Colorado*	\$10.35	N
Delaware*	\$8.78	N
Maine*	\$9.96	N
Massachusetts*	\$12.45	N

Source: LFC survey / GAO report*

the wide variation in contracts and assessments does not enable true ‘apples to apples’ cost comparisons. State standards and state assessments used to measure progress on standards vary. Cost comparisons are difficult as state assessment contracts differ along many characteristics including:

- Number of students tested
- Term of the contract
- Types of test questions
- Percent of items that are extended response
- How answers are scored
- Timeframe for results to be delivered
- Foreign language versions of test
- Ownership of test questions
- Commissioned versus permissioned items
- Number of test forms
- Use of computer-based testing
- Number of items written and released
- Number of research, statistical, or alignment

studies necessary

- Extent to which test questions are released to the public.

Extended response questions cost much more to develop and score. PED reported that, according to Pearson, extended response questions account for 40-50 percent of the cost of the NMSBA. This analysis of various states’ estimated costs per assessment reflects that use of extended question types is one of, if not the major cost driver. In addition, a 2003 Government Accountability Office (GAO¹) report found states with extended response in their assessments had higher costs than states using multiple choice only assessments. The Government Accountability Office (GAO) estimated that states would spend from \$1.9 billion to \$5.3 billion

¹ May 2003, Title I-Characteristics of Test Will Influence Expenses; Information Sharing May Help State Realize Efficiencies, GAO-03-389

between 2002 and 2008 to implement NCLB-mandated tests. If all states used tests with only multiple-choice questions, GAO estimated the total state expenditures would be about \$1.9 billion. If all states used tests with a mixture of multiple choice questions and a limited number of extended response questions, GAO estimated spending to be about \$5.3 billion. For example, Massachusetts, which uses open-ended questions in its tests, incurred a cost of about \$7 to score each test in 2002. In Virginia and North Carolina, which used mostly multiple-choice tests in 2002, the scoring cost was less than \$1 per test.

Benchmark federal funding levels would cover a larger percentage of estimated expenditures if PED chose to use a greater proportion of multiple choice questions. The GAO estimated that federal funds would cover 155 percent of New Mexico's estimated expenditures if it chose to use all multiple-choice questions, but only 81 percent if the state chose to use both multiple choice and extended response questions. The GAO report and a 2002 report from the Education Leaders Council indicate that federal funds provide an amount sufficient to allow for a limited use of extended response questions.

Many states are more judicious in their use of extended response. New Mexico uses multiple-choice, short answer, and extended response questions in all grades tested and in all content areas with 30 percent of the questions being constructed response. Nine states use only multiple-choice for all assessments, reading, math, and science, for all grades tested. Nine other states elected to use constructed response in reading but only multiple-choice questions in math assessments. Of the states that use constructed response, many of them, for example, Delaware, Florida, and Wisconsin, do not use it at all grade levels and in all content areas. Many states that responded to an LFC survey use extended response selectively, most often they do not use it in the early grades as they felt it was less valuable and often not at the high school level as it was even more costly to grade.

Assessment results would be more useful to teachers if they were received in the year the exam was administered. While the NMSBA is not a formative exam and the results are not intended to guide instruction in the sense that they would alter a teacher's focus from one day to the next, the results are useful to identify strengths and weaknesses in curriculum and shift instructional focus or resources, such as instructional coaches or professional development, to address the weaknesses. Other states report that use of results by teachers increased when they got the results back in the year the test was taken. Receiving the results sooner would also provide other stakeholders with a better opportunity to be informed.

The LFC's evaluation of Rio Rancho Public Schools noted that the early NMSBA testing window limits the number of instructional days available before the test. For spring 2009, PED changed the NMSBA test window (3/24/09 - 4/24/09) to be about three and half weeks later than in 2008 to allow for more instructional time before the test, however this also means test results will be received later. Many districts have expressed dissatisfaction with timeliness of results; however, the delivery of results was improved last year. PED delays in 2006 led to districts not receiving results until after November 2006, but test results for the most recently administered NMSBA were delivered to districts around June 24th, 2008.

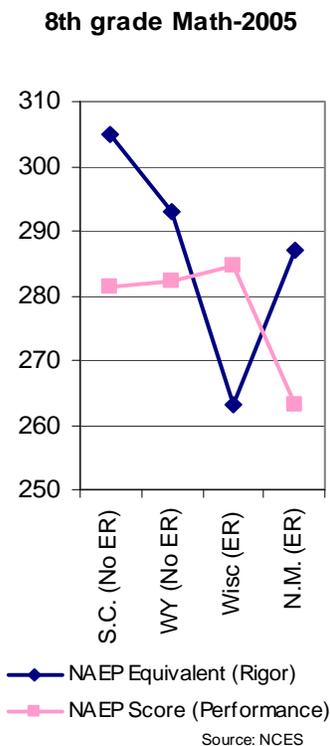
Multiple choice results could be returned before extended response results. Hand scoring of extended response questions is much more time consuming and labor intensive than machine scoring of multiple choice questions. Last year, South Carolina enacted legislation requiring assessments other than writing to be multiple choice. To ensure timely results, this legislation

also requires multiple choice tests to be administered as close to the end of the school year as possible and the writing assessment to be administered earlier in the school year. States using multiple choice only formats typically receive their results sooner than states extensively using extended response. For example, Utah receives raw scores within 24 hours and scaled scores in about 4 days. Other states that use extended response get results back in a time-frame that is comparable. Vermont, for example, gets results back in about 10 weeks. In New Mexico, the current contract requires the contractor to return test results within 60 days of receiving test materials from all districts.

The current NMSBA co-mingles multiple choice questions with constructed response questions for the various competencies. Current test design does not allow for multiple choice scores to be returned separately which could provide districts with more immediate results. To provide quicker feedback, PED could develop future assessments that allow for multiple choice results to be returned before constructed response results.

The reliability of extended response results is a concern. Extended response questions are designed to test students' ability to respond to a prompt. Following a strict scoring rubric helps to maintain an essential level of inter-rater reliability but firm adherence to the rubric can diminish the value of extended response in accurately measuring depth of knowledge. Some districts and educators have expressed doubts about the value and effectiveness of extended response and have raised concerns about the accuracy and reliability of grading extended response. In response to an LFC survey of New Mexico education professionals, five of twelve respondents felt that the scoring of extended response is not reliable, one felt the scoring is reliable, and six were not sure about reliability. Teacher awareness of how rubrics are followed

can lead to measurement of a student's understanding of a rubric rather than their understanding of a standard. In addition to their tremendous costs in terms of time and money, psychometricians from other states question the reliability of scoring and found that extended response questions do not contribute as much information about students' understanding as the tests would indicate.



Extended response often increases the rigor of a state assessment, however there are exceptions.

A rigorous test provides a more accurate measurement of students' performance. Two conventional ways to measure state assessment rigor involve comparisons to the National Assessment of Educational Progress (NAEP), which is considered a very rigorous exam. The first method measures the gap between the percentages of students achieving proficiency on the NAEP compared to the percentages achieving proficiency on the state assessment. A second method involves statistical procedures to estimate a 'NAEP score equivalent', which approximates the NAEP score that a student would have to reach in order to be able to meet proficiency or above on the state assessment. The NAEP score equivalent is an estimated NAEP score that would approximate proficiency on a particular state assessment. A higher 'NAEP score equivalent' indicates a more rigorous state assessment. These measures of rigor demonstrate that, in general,

indicates a more rigorous state assessment. These measures of rigor demonstrate that, in general,

constructed response questions increase rigor and the NMSBA is indeed a rigorous assessment. However, these measures also indicate that use of constructed response does not guarantee a rigorous exam and it is possible to have a well designed assessment with less reliance on constructed response questions and still maintain rigor. For example, South Carolina's assessment uses less extensive constructed response in math and both their 4th and 8th grade assessments have more rigor than New Mexico as measured by NAEP equivalents. California uses only multiple-choice in NCLB assessments and the proficiency gaps are smaller than New Mexico's for both 4th and 8th grade reading. *Measurements of state assessment rigor, such as NAEP equivalent scores and NAEP proficiency gaps, indicate it is possible to have a well designed assessment with less reliance on extended response questions.*

The state is near the bottom in terms of student performance as measured by National Assessment of Educational Progress (NAEP) scores. The NAEP is a national assessment that serves as the common metric of student performance across all states. While the NMSBA is indeed one of the most rigorous and accurate assessments as measured by NAEP comparisons, the state is near the bottom in terms of student performance as measured by actual NAEP scores. For the 2007 NAEP, New Mexico was tied for last in 4th grade math, 45th in 4th grade reading, 48th in 8th grade math, and 49th in 8th grade reading. Student achievement is identified as a state priority and resources could be reallocated from activities that measure student achievement to those that improve student achievement. *There does not appear to be a relationship between the rigor of a state assessment and performance on the NAEP.*

Rigor can be maintained at less expense. New Mexico can continue to meet federal requirements with less use of constructed response. A substantial change in assessments would require a re-submittal of New Mexico's assessment system, which was approved in June 2008, for the Federal Peer Review process to ensure that the state assessment is addressing higher order thinking. Depending on the extent, assessment changes would likely require amendments to the state's accountability plan as well. States with only multiple choice assessments have received Federal approval of their assessment plans and shown that an assessment program does not need widespread use of constructed response to meet the NCLB requirement to employ strategies that measure higher order thinking skills. If developed carefully, with distractor analysis and other item development techniques, high quality multiple-choice questions can be rigorous and measure higher order thinking skills. New Mexico can address the 'depth of knowledge' issue with a reduced use of constructed response.

Other states reduce costs by collaborating on test development. Four New England states, Maine, New Hampshire, Rhode Island and Vermont, are working together to develop tests to meet the requirements of NCLB in a more cost effective manner. The U.S. Department of Education has awarded \$1.78 million to this effort through a competitive grant process. Fourteen states—Arizona, Arkansas, Hawaii, Indiana, Kentucky, Maryland, Massachusetts, Minnesota, New Jersey, North Carolina, Ohio, Pennsylvania, Rhode Island and Washington—joined together to develop a common end-of-course exam in Algebra 2. This is the largest multi-state collaborative assessment effort ever undertaken.

Recommendations

PED should be more selective in the use of constructed response. A reduction in constructed response would greatly reduce costs and allow funds to be reallocated for other important aspects, such as publicly releasing more questions to aid in preparation for the test. A reduction in constructed response could increase the timeliness of results and provide information for schools to make end of year adjustments. PED should develop assessment contracts to make test results from constructed response questions available before the next school year and results from multiple choice questions available in the year the test is administered. A move to a greater use of multiple-choice can be done without an unacceptable reduction in rigor. PED should consider reducing constructed response for math assessments, and focus the use of constructed response in grades 6 through 8 and in the reading and writing content areas. PED could also consider eliminating science assessments when not required, perhaps administering it only in grades 4,7, and 11.

The Request for Proposal (RFP) should allow for variable pricing options. PED should write the RFP for the next NMSBA contract to direct bidders to submit proposals with various levels of constructed response and an estimated time-frame for delivery of results associated with the various levels of constructed response. Bidders should submit a proposal to reduce constructed response from 30 percent to 10 percent in all content areas and in all grades. A second proposal should eliminate constructed response from all math assessments and a third proposal should eliminate constructed response from math and from elementary and high school grades in reading. The RFP should allow the bidder to offer variable pricing so that PED can decide if additional deliverables are worth known additional costs. PED staff indicate they intend to draft the RFP to allow for variable pricing. If PED values additional deliverables or services not necessary to meet statutory requirements, then PED may want to arrange the contract terms so that payments for these services are delayed until FY11 or FY12.

PED should collaborate with other states in the development of certain assessments to reduce costs. New Mexico could realize efficiencies by sharing high development costs associated with low-use tests, such as NMELPA or NMAPA, with our neighboring states.

PED should consider developing a contract so that New Mexico does not have to wait for all districts to get their materials to the contractor before the sixty day clock starts. Some states identify 'priority' districts which are to be graded first. These districts are often representative of overall state performance and by grading these first, the state gets an idea of potential overall results. Nevada uses a staggered reporting system where the contractor scores and delivers results for districts that get their materials in early and a second round of results for districts that get their materials to the contractor later.

APPROPRIATIONS TO PED FOR ASSESSMENT HAVE BEEN SUBSTANTIAL.

Table 2. Federal Funds

FFY02	\$4,474,218
FFY03	\$4,543,618
FFY04	\$4,568,879
FFY05	\$4,595,624
FFY06	\$4,595,624
FFY07	\$4,628,375
FFY08	\$4,581,141
TOTAL	\$31,987,479

Source: LESC/PED

PED received about \$32 million in federal appropriations from Federal Fiscal Year (FFY) 02 to FFY08 for the development and implementation of NCLB mandated assessments.

Federal funds are used for standards-based assessments in reading or language arts, mathematics, and science. When a state has developed the required standards and assessments, any additional funds may be used to develop content and achievement standards and to ensure the reliability and validity of state assessments. In testimony to the LESC during the 2005 Interim, PED reported that the federal appropriations were initially used exclusively for the development of grades 3-9 standards based assessments, but for FY05, the federal funds were also

used to supplement the costs of developing the grade 11 assessment and for related costs such as the Technical Advisory Committee contracts and an Adequate Yearly Progress (AYP) calculator.

The New Mexico Legislature has appropriated over \$21.1 million to PED for development and implementation of state assessments for fiscal years FY05 through FY09.

For FY08, the Legislature appropriated \$2.0 million from the education “Lock Box” to PED: \$500,000 to develop a federally required alternative assessment; and \$1.5 million to develop a new 11th grade assessment. In 2008, the Legislature reauthorized the use of the \$1.5 million for expenditure through FY09 and appropriated an additional \$4.0 million to PED for assessments, resulting in a total of \$5.5 million available for expenditure in FY09. Funds from the “Lock Box” are to be spent only for the implementation and maintenance of education reform. *The Legislature appropriated \$1.5 million for the 11th grade assessment; however PED intends to use these funds for an information technology project, the electronic student management system.*

Table 3. State Funds for PED's Assessment Program

Year	Amount	Description
FY05	\$810,000	Sect 6 supplemental for over-obligation of NCLB testing contract
FY05	\$350,000	Sect 5 Special nonrecurring: state level assessment costs
FY05	\$385,500	Sect 5 Special nonrecurring: student assessment
FY06	\$7,000,000	Sect 5 Special nonrecurring: assessment and development from Lock Box
FY07	\$6,600,000	Sect 5 Special nonrecurring: assessment and development from Lock Box
FY08	\$1,500,000	Sect 5 Special nonrecurring: 11th grade exam from Lock Box (ext thru FY09)
FY08	\$500,000	Sect 5 Special nonrecurring: SE alternative assessment from Lock Box
FY09	\$4,000,000	Sect 5 Special nonrecurring: assessment and development from Lock Box
TOTAL	\$21,145,500	

Source: LFC files

Significant funds are distributed through the State Equalization Guarantee (SEG) for district expenses related to assessment. From FY01 to FY09, the Legislature appropriated more than \$7.8 million in additional recurring dollars through the Public School Funding Formula to provide for district assessment costs. Beginning in FY06, significant appropriations were also provided to PED as non-recurring special appropriations from Section 5 of the General Appropriation Act (GAA).

Table 4. State Funds to Districts

Year	Amount	Description
FY01	\$1,000,000	Above line Program Cost (recurring)
FY02	\$600,000	Above line Program Cost (recurring amount of 1.6 million)
FY04	\$2,000,000	Above line Program Cost (recurring amount of 3.6 million)
FY05	\$1,800,000	Above line Program Cost (recurring amount of 5.4 million)
FY06	\$2,000,000	Above line Program Cost (recurring amount of 7.4 million)
FY07	\$0	(recurring amount of 7.4 million)
FY08	\$0	(recurring amount of 7.4 million)
FY09	\$464,800	Above line Program Cost (recurring amount of \$7,864,800)
TOTAL	\$7,864,800	

Source: LFC files

District costs for the NMSBA are a flat \$30 per assessment for each district and allocating funds through SEG provides some districts with more for assessment and some with less because of significant variation among districts in funding units per member.

Total costs for the NMSBA contract from FY06 to FY09 are about \$31.1 million. Expenses

Table 5: NMSBA Contract Costs

Year	PED	Districts*
FY06	\$1,506,279	\$0
FY07	\$5,947,333	\$1,112,902
FY08	\$4,391,275	\$6,597,059
FY09	\$4,835,743	\$6,704,038
TOTAL	\$16,680,630	\$14,413,999

* - \$30 per student for NMSBA / \$25 per student for NMHSCE

Source: PED PSC #06-924-P527-0143

for the NMSBA are shared between the PED and the districts. The current contract requires PED to pay for initial test development. The contract assigns PED a total cost from FY06 through FY09 of nearly \$16.7 million. School districts pay for the production of NMSBA materials, scoring, and reporting of assessment results.

School districts pay \$30 per NMSBA ordered. Total contract costs assigned to districts are about \$14.4 million.

Appropriations exceed expenses documented for the NMSBA, NMELPA, NMHSCE, and NMAPA assessments. As shown on Table 6, from FFY05 to FFY08, federal appropriations to PED for the development and implementation of assessments were over \$18 million. From

Table 6. Assessment Funds and Contractual Expenses

Year	Federal Funds to PED	Year	State Appropriations to PED	PED assessment contracts (FY06-FY09)*
FFY05	\$4,595,624	FY06	\$7,000,000	\$7,506,199
FFY06	\$4,595,624	FY07	\$6,600,000	\$8,093,526
FFY07	\$4,628,375	FY08	\$2,000,000	\$6,007,976
FFY08	\$4,581,141	FY09	\$4,000,000	\$6,858,455
Total	\$18,400,764	Total	\$19,600,000	\$28,466,156

* - Expenses only for the NMSBA, NMELPA, NMHSCE, NMAPA

Source: Expenses from PSC #06-924-P527-0144, PSC #06-924-P527-0157, PSC #06-924-P527-0143

FFY06 through FY09, the Legislature appropriated \$19.6 million to PED for development and implementation of state assessments. Over roughly the same period of time, PED expenses for the main assessment contracts were about \$28.5 million or about

75 percent of the available state and federal funding. The Assessment and Accountability Division has incurred expenses beyond those shown in Table 6. For example, PED pays other contractors to make AYP determinations. PED provided information on January 9, 2009 describing these additional expenditures; however, this information has not yet been fully verified by LFC staff.

Recommendations

PED must provide the LFC with a detailed report and remaining documentation on assessment division expenditures for FY06 through FY09. PED should completely account for how all assessment appropriations have been used prior to receiving any additional special appropriations or contractual service funds for this purpose.

The federal appropriation amounts should be used as the target for the NMSBA contract cost and help to establish a more reasonable balance between the use of constructed response and multiple choice questions.

The Legislature should consider moving funds for assessment out of the SEG and providing them to districts as a categorical appropriation to make assessment funding more equitable. This idea has been discussed with LESC in the 2008 session and during the interim.

PED's INADEQUATE CONTRACT MANAGEMENT RESULTS IN EXCESSIVE COSTS.

Tracking of contract deliverables and monitoring of contractor performance needs improvement. PED's contract monitoring documents do not track all of the deliverables listed in the scope of work, nor do they record contractor performance on the contract performance measures. PED tracks occurrences when districts notify PED that test materials have not been received but over 30 of these issues had no resolution listed. This suggests either a level of contractor performance that does not meet expectations or inadequate monitoring of issue resolution by PED.

Improved tracking of deliverables by PED would help to ensure satisfactory contractor performance. For example, the contract requires a "Process audit" to be conducted within six months after the contract is awarded. The contract was signed on February 2006 and the process audit was initiated in January 2008 and completed in April 2008. If PED had a more effective tracking system in place, they could have notified the contractor that the process audit was overdue. The contract appropriately requires the contractor to produce time and expenditure records for services provided; however PED reported that they do not audit the contractor's invoices. PED provided copies of warrants for five of the seven invoices submitted thus far.

PED is paying for items not received. Documents provided by PED indicated that several deliverables specified in the contract have not been delivered. These include demonstration books, classroom level reports, a released item report, a district level item plot report, a school level item plot report, error summary reports, cognitive labs, and PDF versions of management reports. The management plan refers to a detailed project schedule to be provided to PED in PDF form; this schedule was requested by LFC but never provided to PED by the contractor.

PED has not sought liquidated damages for documented poor performance. Documents provided by PED show a calculation of liquidated damages of \$55,000 for late delivery of materials; however, these damages were not collected. PED indicated liquidated damages were not pursued because the contractor resolved the problem by resending districts test booklets without errors; however the contract would still allow for a collection of liquidated damages.

The contract contains vague language about PED responsibilities which can impact the ability to hold the contractor accountable. The contractor can not be held accountable when "acts or omissions of the agency result in the Contractor's inability to provide services in a timely manner" or "if agency staff has not met their portion of the schedule." PED responsibilities are not clearly specified in the contract. The contract requires "Annual written program plans, signed by both parties", a "Comprehensive management plan", and an "Implementation plan". PED provided a project management plan which is intended to function as all of the contractually required plans; however, this document did not contain well organized or clearly delineated agency responsibilities with specific timelines attached. According to PED, agency responsibilities are ostensibly to be found throughout this 'project management plan'; however this plan indicates that relevant documentation is to be managed and maintained by Pearson rather than PED. This makes a proper evaluation of contractor performance unfeasible as the contractor is responsible for keeping some of the documentation reflective of their own performance. For example, a key performance measure is timeliness of results which are due 60

days after Pearson receives all districts materials. PED did not provide documentation of when each district provided the contractor with test materials but reported that the contractor keeps a record of when test materials are delivered to and returned from districts.

In FY08, procurement violations over \$1.5 million occurred in the three main assessment contracts often related to PED requesting services outside the scope of the contracts. On August 10, 2007, PED sent a letter notifying the State Purchasing agent of procurement violations of \$953,933 in the NMSBA contract stating that PED “was unaware that the request was outside the contract scope.” Procurement violations in the NMELPA contract with Pearson occurred because PED under-estimated the number of test booklets needed. Pearson manufactured and delivered materials to cover the shortage just before administration of the NMELPA at a cost of \$90,586. There were procurement violations for \$542,190 regarding the contract with AIR for the NMAPA. A PED Office of Inspector General (OIG) report determined that an amount billed by AIR exceeded what would be considered ‘fair and reasonable’ and found that contract deliverables were not completed. An amendment extending the contract with AIR was signed after the issuance of the OIG report. The OIG report recommended PED assessment division staff training in contract management but PED did not report that training had occurred. This evaluation concurs with the finding in the OIG report which states the Assessment Division needs to develop “contract management expertise that extends beyond policy and procedures to the sphere of effective monitoring and managing of contractor performance.”

For the NMAPA, PED selected a contractor with a bid twice as high, or over \$1.2 million more, than the lowest bidder. PED’s bid process for the NMAPA contract did not produce satisfactory results. The OIG report on the NMAPA contract noted that PED selected AIR, the highest bidder, because PED thought the superior quality of AIR products and services justified the higher price. The next highest bidder was 20 percent lower and the third bidder was 50 percent lower than AIR, which would have amounted to over \$1.2 million in reduced costs. The report also states former PED employees acknowledged that AIR’s performance has been disappointing. AIR also conducted the study on the current public school funding formula.

PED recently amended the NMSBA contract which amounted to nearly \$425,000 in additional FY09 costs. On October 17, 2008, Governor Richardson issued a press release directing agencies to reduce costs and contract expenditures. PED submitted and DFA approved on 10/20/2008, an amendment to the NMSBA contract. The amendment requires the contractor to develop the Spanish 11th grade SBA in science and writing as the original contract only required 11th grade Spanish SBA in math and reading. The August 2007 letter from PED to State Purchasing states that the 11th grade Spanish Science and Writing assessments have been developed with work beginning in 2006.

PED has made some recent improvements to contracts. Amendments to the NMELPA and NMAPA contracts in 2008 have provided greater detail in specifying requirements for deliverables. For example, an amendment to the NMAPA contract clarified requirements for the test materials ordering system and an amendment to the NMELPA contract clarified data reporting and labeling requirements. PED’s recent efforts to increase specificity in deliverables should improve results.

Recommendations

PED must improve their management of contracts. PED should assume more responsibility for maintaining documentation reflective of deliverables and contractor performance. PED needs further training in development of effective RFP and contract terms, negotiation techniques, and performance monitoring. For every contract, PED should develop a complete list of deliverables to be tracked and a contract abstract which clearly summarizes the services to be provided in the scope of work to help avoid further procurement violations resulting from requests for services that are not in the scope of work.

Deliverables need specific measures of performance. Contracts should include unambiguous measures of quality for each deliverable described in the scope of work which would enable a more complete evaluation of contractor performance. In order to better monitor contractor performance and competitiveness of pricing, the contractor should be required to keep and provide records of incurred costs for major deliverables.

Details about agency responsibilities and associated timelines need to be more clearly defined in contracts. Agency duties must be organized better in planning documents, particularly where agency responsibilities impact contractor accountability. Specifying agency and contractor responsibilities and detailing deliverables ensures the state receives exactly what is desired. Omitted details are potential points of contention and items or services for which the contractor will charge extra.

AGENCY RESPONSES

STATE OF NEW MEXICO
PUBLIC EDUCATION DEPARTMENT
300 DON GASPAR
SANTA FE, NEW MEXICO 87501-2786
Telephone (505) 827-5800
www.ped.state.nm.us

DR. VERONICA C. GARCÍA
SECRETARY OF EDUCATION

BILL RICHARDSON
Governor

January 12, 2009

MEMORANDUM

TO: David Abbey, Director, Legislative Finance Committee

FROM: Veronica C. García
Secretary of Education

RE: PED RESPONSE TO PROGRAM EVALUATION OF ASSESSMENT PROGRAM

The January 6, 2009, draft report, *Public Education Department Review of Assessment Program*, by the Legislative Finance Committee has been carefully reviewed, and I am submitting the PED's response to the findings in the draft report. My staff has been working diligently with yours to clarify questions and provide program, funding, and expenditure information necessary to complete the evaluation of the New Mexico Statewide Assessment Program (NMSAP).

I am submitting this response to clarify information and conclusions presented in the draft report. We recently submitted more detailed funding and expenditure information, at your staff's request. This information will need to be added to the report to provide a complete and accurate picture of the cost of the NMSAP. The PED response also provides information regarding questions about contract deliverables and contract management activities, and it is essential to understanding what has occurred. I am confident that this information will receive careful consideration so that our legislators will be able to derive accurate conclusions about our contracts.

I want to credit the Assessment and Accountability staff for making significant improvements in contract writing, monitoring deliverables, establishing procedures to address contractor performance issues, and following the State Procurement Code in the last year and a half. I trust that the recommendations in the final report will consider these improvements.

In the last two years, our state has received national recognition for its high quality student assessment program, and I look forward to working with your staff to advance that legacy into the future.

Comparison of Assessment Costs to Other States

The report provides a comparison of assessment costs among a sample of states. In Table one, page five, New Mexico is identified as one of the states with the highest costs. The information presented in Table one is based on two sources of information, an LFC survey completed recently and a 2003 report by the Government Accounting Office (GAO). Cost information on New Mexico is from the recent survey and costs from a number of other states with reportedly low costs are from the 2003 report. It is difficult to draw conclusions about the relative cost of New Mexico's assessment, which was recently obtained from the LFC survey, with apparently low cost states reported by the GAO six years ago.

Extended Response Questions

Page five, paragraph three, states that extended response questions cost much more to develop and score, and that 40 to 50 percent of the cost of the (Standards-based Assessment) SBA is attributed to extended response items. The percentage statistic reflects the cost of item development, production of test materials, and scoring over the four-year term of the existing contract. New Mexico has made a substantial investment in funding and effort for the development of extended response items and scoring rubrics. New Mexico now owns those items and rubrics, and they can be transferred to a new contractor, reducing the cost of using constructed response items in the future. Since the merger of our initial vendor, Harcourt Assessment, Inc., with NCS Pearson, the scoring capacity of the new vendor has been expanded. With this expanded capacity, PED has authorized the application of "distributive scoring," allowing constructed and extended responses to be scored more efficiently using electronic technology.

Page seven, paragraph three, states that the "reliability of extended response results is a concern." The statement raises a question. However, the report does not provide research-based statistics regarding the comparable reliability of multiple choice and extended response items to support a conclusion.

Funding and Expenditures

Table five, page 11, reports PED and school district expenditures from the second contract with Harcourt for the SBA (term FY06 through FY09). However, Table five lacks certain information on expenditures for the first SBA contract that was effective during FY05 through FY07 and overlapped with the second SBA contract in FY07 (photocopy of contract provided to the LFC). Further, school district expenditures in FY09 from SBA Contract Amendment 1 are not reported in table four (photocopy of contract amendment provided to the LFC). Inclusion of this information is necessary to provide a complete report of SBA contract expenditures.

Table six, page 11, reports federal and state funding and PED expenditures for the NMSAP for the FY06-FY09 period. Table six does not include numerous expenditures for consultation with national assessment experts regarding technical issues, special studies necessary to demonstrate the validity of assessments and alignment with content standards, and professional development activities to prepare test data and instruct school districts on the use of test score information for planning instructional interventions. Further, table six does not include expenditures for item and development, production of test materials, and reporting of test results reported in FY08. These expenditures were necessary to meet U.S. Department of Education requirements for the technical quality and administration of assessments.

Contract Deliverables

Page 13, paragraph one, states that the PED does not track all contract deliverables and it cites deliverables listed without notations as evidence of inattention to deliverables and their quality. Paragraph three further states that the PED is paying for items not received. Items included in the list are summarized in paragraph three of page 13 of the LFC report. Many of the deliverables listed were in development and not yet delivered at the time they were inventoried. The following information is provided to clarify the status of the deliverables specified on page 13 of the report.

- Demonstration booklets were for grade 11 and were not needed by the PED.
- Classroom-level reports were not required because of a lack of data relating student test scores to teachers; difficulties in attributing student test performance to a single teacher, particularly for middle and high school students; and recognition of the potential risks of disclosing confidential information about students and teachers. The contractor informed the PED that reporting at the classroom level was not budgeted and no credit was available for removing this deliverable from the Scope of Work.
- Released Item Reports have been delivered to the PED as required in November 2005 and November 2006. No test items were released from the 2007 administration of the SBA due to the removal of the third core of items from the Scope of Work. A reduction in costs was effected in Contract Amendment 1.
- District-Level Item Plot Reports were delivered in October 2008.
- School-Level Item Plot Reports were delivered in October 2008.
- Cognitive Labs were substituted with another method for field testing assessment items. The change was cost neutral.
- PDF versions of management reports were not needed by the PED. The contractor provided reports in hardcopy format in three-ring binders.
- Financial audit is not a deliverable as such. It is related to the standard requirement in the State Professional Services Contract Template that contractors must retain detailed records of time and expenditures for up to three years from the date of termination of the contract. The PED has not requested an audit of the contractor.

Page 13, paragraph four, states that the PED has not sought liquidated damages for documented poor performance. The instance cited notes that the contractor provided a remedy that satisfied the needs of the school districts. The assessment contracts include penalties for contractor failures to deliver as specified. Recently written contracts more often contain penalty clauses. The Assessment and Accountability Division of the PED is prompt in notifying contractors of poor performance in the form of a cure notice. The Division informs the contractor of the performance issue and requires a response that will meet the conditions of the contract. Contractors are responding with satisfactory remedies. The Division is working diligently to maintain effective long-term working relationships with contractors that are based on trust. The Division will invoke penalties when remedies are not satisfactory and trust as been violated.

Procurement Violations

Page 14, paragraph two, of the report states, “In FY08, procurement violations over \$1.5 million occurred in the three main assessment contracts often related to PED requesting services outside the scope of the contracts.” The PED identified and reported the procurement violations in FY08. The SBA and New Mexico High School Competency Exam (NMHSCE) violations occurred during FY06 and FY07. The violations associated with the New Mexico English Language Proficiency Assessment (NMELPA) occurred during FY06 and FY07. Violations associated with the New Mexico Alternate Performance Assessment (NMAPA) occurred in FY07 and FY08.

Page two, paragraph five, suggests that the procurement violations occurred due to lapses in contract monitoring and tracking of deliverables. Though these activities are essential to managing contracts, the procurement violations are due to a number of causes unrelated to monitoring and tracking. The following paragraphs provide a brief summary of causes of the procurement violations.

- The first administration of the fall 2006 NMHSCE under the new contract did not provide sufficient time to score and report student pass/fail results in time for districts to process winter graduates. The PED requested that the contractor provide an expedited reporting and delivery of student-level pass/fail information from the fall 2006 test in order to meet winter graduation.
- During May 2006 SBA review to prepare data for determination of adequate yearly progress (AYP) for schools and districts, the PED processed a large volume of data changes, delayed delivery of data to the contractor, and required the contractor to implement an additional cycle of scoring and reporting to produce and distribute school, district, and state summary reports on 2006 NMSBA results.
- The PED is required to implement a statewide standards-based assessment in science in one grade in the grade 10 through 12 span by the 2007-2008 school year [20 U.S. Code Section 6311 (a)(2)] and (Section 22-2C-4(E), NMSA 1978). In summer 2006, the PED requested that the contractor develop a field test version of the assessment and to conduct a field test in April 2007.

- New Mexico statute specifies that the PED will develop and implement a standards-based academic performance writing assessment in Spanish (Section 22-2C-4(B, 2), NMSA 1978) for grades three through nine. Work on the Grade 11 Spanish writing assessment began in 2006.
- The PED underestimated the number of test booklets and answer sheets needed for the first administration of the NMELPA, which number proved insufficient to fill school district orders. The PED requested that the contractor manufacture additional test booklets and answer sheets outside the scope of the contract.
- Under PED's directive, the districts administered two out of four sections of the NMELPA to kindergarten students in January 2007. However, this directive was overridden by the U.S. Department of Education. In order to comply with this directive, it was necessary to retest 317 kindergarten students in all four areas in April 2007. PED requested that the contractor print test booklets and answer documents, and distribute the materials to schools for retesting.
- AIR conducted an alignment study to determine if the performance tasks comprising the spring 2007 NMAPA met state and federal standards for test quality. A number of tasks failed to meet technical requirements and needed to be replaced with new items to be administered in the spring 2008. Under PED's request, AIR designed and wrote tasks and other necessary materials.
- AIR conducted meetings with a select group of New Mexico special education teachers to review new NMAPA performance tasks and to identify issues with content and cultural bias. AIR paid for the meeting facilities, and the costs of participants' travel, hotel rooms, and meals. AIR staff traveled from Washington, D.C. to facilitate and record the meetings.
- AIR conducted meetings with a select group of special education teachers from across New Mexico and the PED to develop instructional standards. AIR prepared materials necessary to conduct the meetings. AIR paid for the meeting facilities, and the costs of participants' travel, hotel rooms, and meals. AIR staff traveled from Washington DC to facilitate and record the meetings.
- AIR developed manuals necessary to administer and score the spring 2007 performance tasks in language arts, math, and science that were released by the PED. This work was completed on August 31, 2007.

Contract Expertise

Dr. Tom Dauphinee, Deputy Director of Assessment and Accountability at the PED, assumed responsibility for managing contracts in January 2007. He identified and reported the said procurement violations. No procurement violations have occurred since the summer of 2007. Dr. Dauphinee participated in the training class on contracting offered by the Contract Management Bureau of GSD. He also purchased and studied materials on government contracts. He is investigating resources that can provide additional training on the State Procurement Code, the RFP process, and contract negotiation, management, and writing.

Dr. Dauphinee has prepared many contract amendments for the development of student assessments and ancillary professional services since January 2007. A concerted effort has been devoted to providing detailed specifications for deliverables in contracts, including contract language for fiscal penalties for contractor failures to meet requirements, seeking regular feedback from Assessment and Accountability staff regarding the quality of contractor deliverables, establishing a cure notice process and consistently communicating with contractors regarding failures and necessary corrective actions, preparing accurate cost estimates for contract deliverables, adhering to State Procurement Code requirements regarding authorized procurement procedures, providing training to Assessment and Accountability staff regarding authorized procurement procedures, and requesting funding for the PED and school districts through the legislative appropriation process.

Contract Amendment

Page 14, paragraph four, states that the PED recently amended the SBA contract, increasing state costs at a time when Governor Richardson directed agencies to reduce costs. The contract amendment was in preparation for a number of months, and it was signed in September 2008, preceding Governor Richardson's press release of October 17, 2008. Further, the amendment authorized the contractor to complete the development of the Grade 11 Spanish SBA in science, as required by NCLB [20 U.S. Code Section 6311 (a)(2)] and (Section 22-2C-4(E), NMSA 1978); and writing, as required by (Section 22-2C-4(B, 2), NMSA 1978). NCLB requires that federal funds be used for standards-based assessments in science. The amendment also provided student test data in a format usable to school districts. Student Record Labels were also provided to affix test scores to student cumulative folders, which had been requested by school districts for several years. Additional funding for school districts to pay for these deliverables was appropriated through the funding formula by the 2008 Legislature, in response to the PED's request.

Thank you.

VCG/TD/rmw

cc: Sen. John Arthur Smith, Chairman, Legislative Finance Committee
Rep. Luciano "Lucky" Varela, Vice-Chairman, Legislative Finance Committee
Rep. Henry "Kiki" Saavedra, Chair, LFC Program Evaluation and Information
Technology Subcommittee
Rep. Rick Miera, Chair, Legislative Education Study Committee
Sen. Cynthia Nava, Vice-Chair, Legislative Education Study Committee
Brian Condit, Chief of Staff, Office of the Governor
Katherine Miller, Secretary, Department of Finance and Administration
Dr. Peter Winograd, Director, Office of Education Accountability
Frances Maestas, Director, Legislative Education Study Committee
Manu Patel, Deputy Director, Legislative Finance Committee
Don Moya, Deputy Secretary, Finance and Operations, PED
Dr. Catherine Cross Maple, Deputy Secretary, Learning and Accountability, PED
Carlos Martinez, Assistant Secretary for Assessment and Accountability, PED
Cynthia Marietta, Director, Administrative Services Division, PED
Willie Brown, General Counsel, PED

STATE OF NEW MEXICO
PUBLIC EDUCATION DEPARTMENT
300 DON GASPAR
SANTA FE, NEW MEXICO 87501-2786
Telephone (505) 827-5800
www.ped.state.nm.us

DR. VERONICA C. GARCÍA
SECRETARY OF EDUCATION

BILL RICHARDSON
Governor

January 13, 2009

MEMORANDUM

TO: David Abbey, Director, Legislative Finance Committee

FROM: Secretary of Education Veronica C. García

RE: ADDENDUM -- PED RESPONSE TO PROGRAM EVALUATION OF ASSESSMENT PROGRAM

During our January 12, 2009, meeting to discuss the PED response to the LFC report, *Public Education Department Review of Assessment Program*, your staff requested additional information on the bid process in 2005 for the New Mexico Alternate Performance Assessment (NMAPA). Specifically, a justification was requested for selecting the American Institutes for Research (AIR), though AIR's bid was more than \$1.2 million higher than the lowest bidder.

I am submitting this Addendum as a response to that request. This information explains the regulatory context of the bid process and the Evaluation Committee's recommendation to select AIR to develop and implement the NMAPA.

Selection of Contractor for the New Mexico Alternate Performance Assessment

In 2005, the State of New Mexico was under a timeline waiver from the U.S. Education Department (USED) to implement an alternate assessment for students with significant cognitive disabilities that met NCLB technical requirements. The waiver provided one year for the State to comply under the potential threat to withhold a significant portion of Title 1 Administrative funding from the PED.

In order to maximize the probability of selecting a successful contractor and receiving USED approval of the NMAPA, the RFP Evaluation Committee emphasized the importance of a strong management component and a strong project plan in the evaluation score points. The committee placed 30% of the 1,000 points possible on the management component, 40% of the points on the project plan, 20% of the points were allocated to the cost component, and the remaining 10% were assigned to the oral presentation.

Three proposals were received. The Evaluation Committee awarded AIR an average score of 291.67 out of 300 on the management component, while the second Offeror received an average score of 217, and the third Offeror received an average score of 116.67. The Committee determined that the corporate experience, corporate references, staff experience, and individual references for AIR were superior with respect to expertise and understanding of assessment of students with significant cognitive disabilities and New Mexico's desired test design. The Committee determined that the second and third ranked Offerors held insufficient experience with the specialized student population to successfully implement the State's test design.

The Evaluation Committee awarded AIR an average score of 388.33 out of 400 on the project plan component. The second and third ranked Offerors were awarded 250 and 116.67 points, respectively. The Committee determined that AIR proposed a test design that was the most relevant and appropriate for New Mexico's population of students with significant cognitive disabilities. The Committee also believed that AIR proposed the most feasible test design for a one year implementation schedule, which the timeline waiver required.

AIR was the lowest ranked Offeror on the cost component, receiving an average score of 82 points out of 200. The second ranked Offeror overall was also second ranked on the cost component, scoring 103 points. The lowest ranked Offeror overall was the highest scorer on the cost component with a perfect score of 200.

AIR was awarded 100 points on the oral presentation, the highest average score. The second ranked Offeror overall had the second ranked score on the oral presentation (71.67 points). The lowest ranked Offeror overall also had the lowest score on the oral presentation (16.67 points).

Based on the evaluation criteria and scoring system, AIR was awarded the highest score overall. Therefore, the Committee recommended that AIR be selected to develop and implement the NMAPA.

Thank you.

cc: Sen. John Arthur Smith, Chairman, Legislative Finance Committee
Rep. Luciano "Lucky" Varela, Vice-Chairman, Legislative Finance Committee
Rep. Henry "Kiki" Saavedra, Chair, LFC Program Evaluation and Information
Technology Subcommittee
Rep. Rick Miera, Chair, Legislative Education Study Committee
Sen. Cynthia Nava, Vice-Chair, Legislative Education Study Committee
Brian Condit, Chief of Staff, Office of the Governor
Katherine Miller, Secretary, Department of Finance and Administration
Dr. Peter Winograd, Director, Office of Education Accountability
Frances Maestas, Director, Legislative Education Study Committee
Manu Patel, Deputy Director, Legislative Finance Committee
Don Moya, Deputy Secretary, Finance and Operations, PED
Dr. Catherine Cross Maple, Deputy Secretary, Learning and Accountability, PED
Carlos Martinez, Assistant Secretary for Assessment and Accountability, PED
Cynthia Marietta, Director, Administrative Services Division, PED
Willie Brown, General Counsel, PED

APPENDIX A: NEW MEXICO STATE ASSESSMENT PROGRAM

NEW MEXICO STATE ASSESSMENT PROGRAM:

- New Mexico Standards Based Assessment (NMSBA) is developed to meet the requirements of the No Child Left Behind Act. The NMSBA is based on New Mexico state standards, which define what students should learn each year, and is administered to all New Mexico students in grades 3-8 and 11. PED contracts with Pearson for the NMSBA with the same contract used for the NMHSCE. PSC#06-924-P527-0143
- New Mexico Alternate Performance Assessment (NMAPA) is an alternative assessment designed for the small number of students who are unable to participate in regular grade-level assessments. NMAPA ensures that all students are included in the statewide assessment program. In May 2005, PED contracted with AIR to develop the NMAPA. PSC#06-924-P527-0157
- New Mexico English Language Proficiency Assessment (NMELPA) measures English language proficiency of new students whose native language is not English. NMELPA helps schools determine whether a student needs English as a second language instruction or can be placed directly into regular classes and establishes an indicator of English language proficiency that facilitates monitoring of student progress. In February 2006, PED contracted with Harcourt Assessment to develop the NMELPA. PSC#06-924-P527-0114
- New Mexico High School Competency Exam (NMHSCE) is administered to all students in Grade 10, and to any students in Grades 11, 12, and 13 who have not passed all subtests. A passing score on all subtests is required in order to receive a high school diploma. PED contracts with Pearson for the NMHSCE with the same contract used for the NMSBA. PSC#06-924-P527-0143
- National Assessment of Educational Progress (NAEP) assessments are administered uniformly across the nation using the same sets of test booklets, so NAEP results serve as a common metric for all states and selected urban districts. NAEP is usually administered at grades 4 and 8 in math and reading and results are based on a representative sample of students in public schools. ETS, Westat, Pearson, and AIR all have contracts for various aspects of the NAEP program.