Teacher and Administrator Preparation in New Mexico

LFC Program Evaluation Presentation to the LESC

Michael Weinberg, Ed.D.

July 11, 2013

Introduction: The Importance of Teacher Preparation in NM

- Half of New Mexico's K-12 teachers prepared at six NM colleges of education
- In FY12, $1.2 billion of district/ charter budgets to salaries and benefits
- Colleges of education account for 11 percent of credit hours and generate $64 million in formula funding
- In FY12, 51% on grade-level in reading, 43% in math

Number of Licensed Teachers by College of Education*

	Licensed Teachers	Percent of Total
CNM	420	1%
Eastern	1,792	4%
Highlands	1,333	3%
NMSU	2,859	6%
UNM	5,368	11%
Western	815	2%
Other**	13,077	28%
Unknown	21,758	46%
Total	**47,422**	

*Not all licensed teachers are actively teaching.
**Includes teachers prepared by private, out-of-state, and other in-state institutions.

Source: LFC Analysis

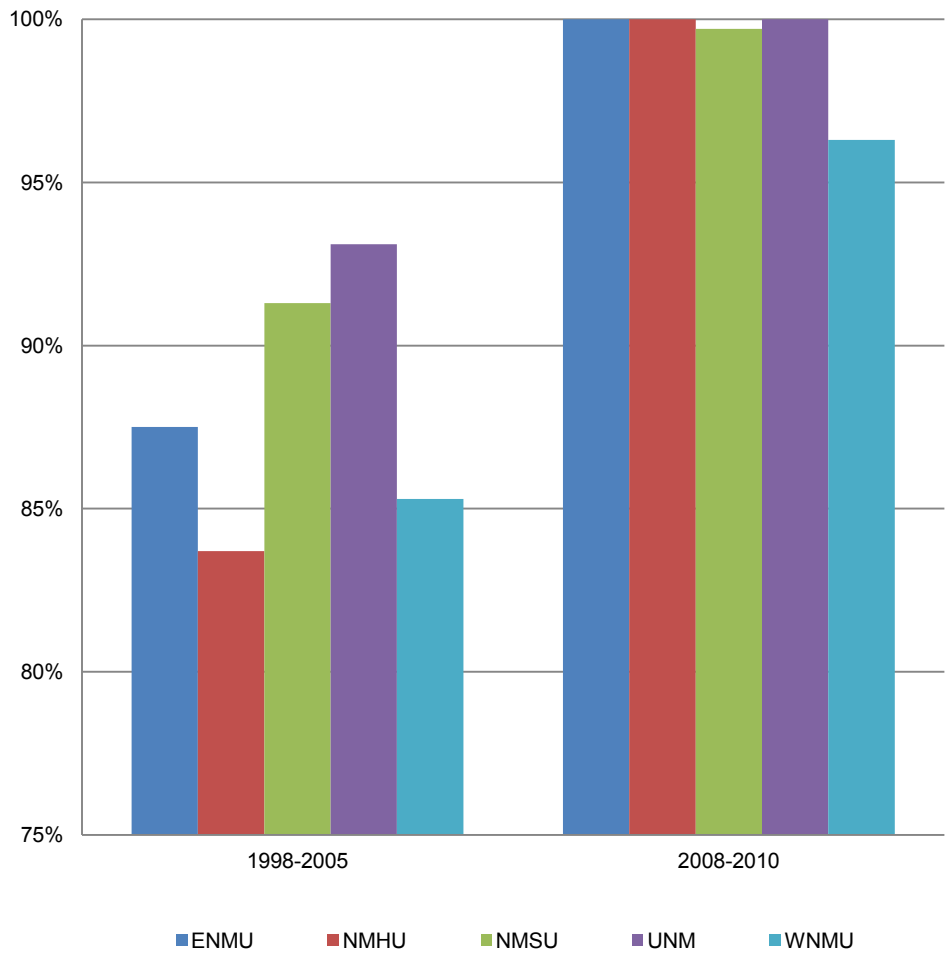
Key Findings

1. Admission requirements and licensure standards perpetuate low student performance
2. PED could better oversee preparation programs
3. Increasing entrance standards, exit standards, and program quality for administrators
4. Improvements to EARS

Finding One: Low teacher admission requirements

- Invested $59 million in salary increases through three-tiered system since 2009
- Basic skills test as entrance requirement—low standard, high pass rates
- CAEP draft accreditation standards raise admissions requirements

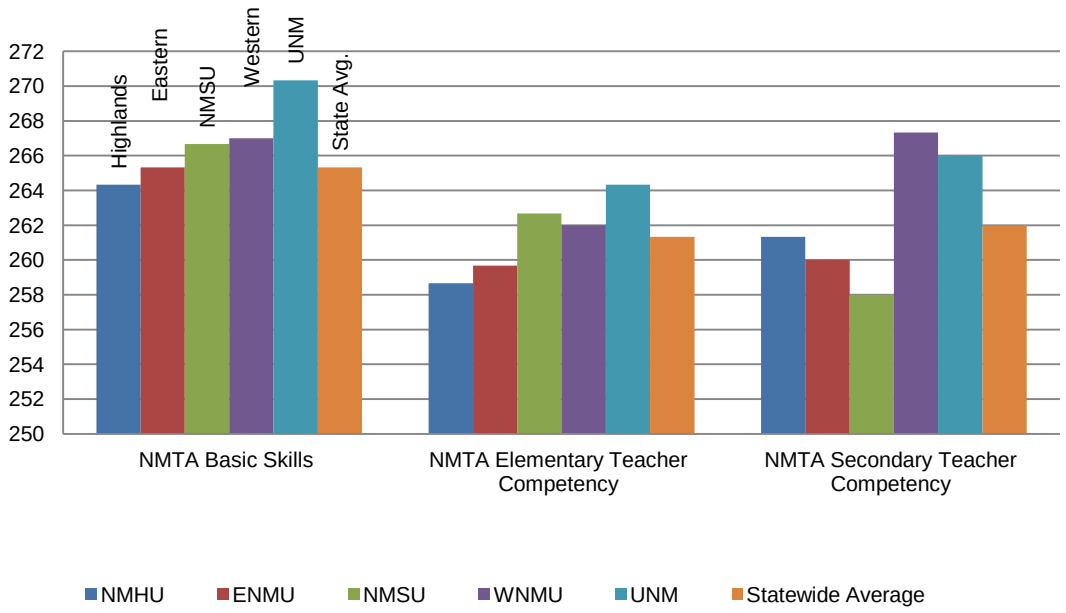
Chart 5. NMTA Basic Skills Pass Rates



Source: LFC Analysis

Finding One: Low Licensure Standards

Chart 8. Traditional Preparation Program NMTA Scores, SY08 - SY10

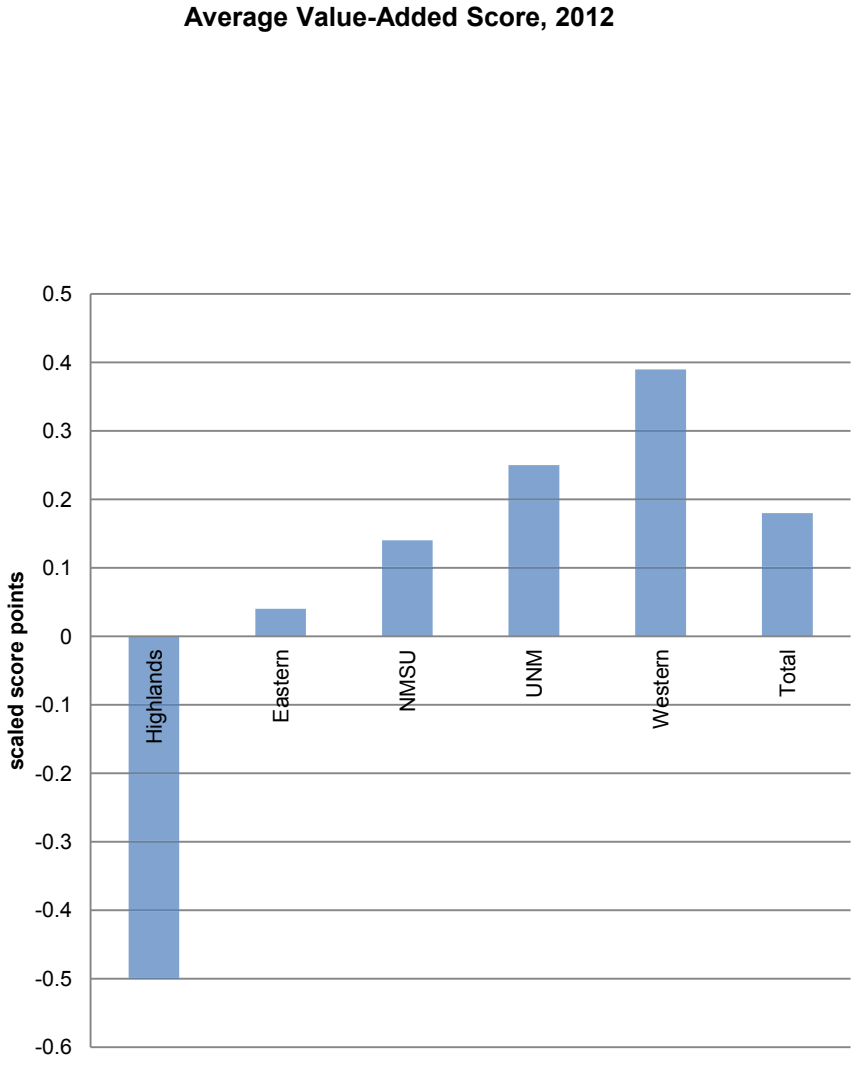


Source: 2011 Title II Report

- Passing scores for NMTA assessments set one standard deviation below average
- Teachers who fail an NMTA at least one time perform lower than those who pass on their first attempt.
- NMTA scores correlate with student performance

Finding Two: Better Oversight

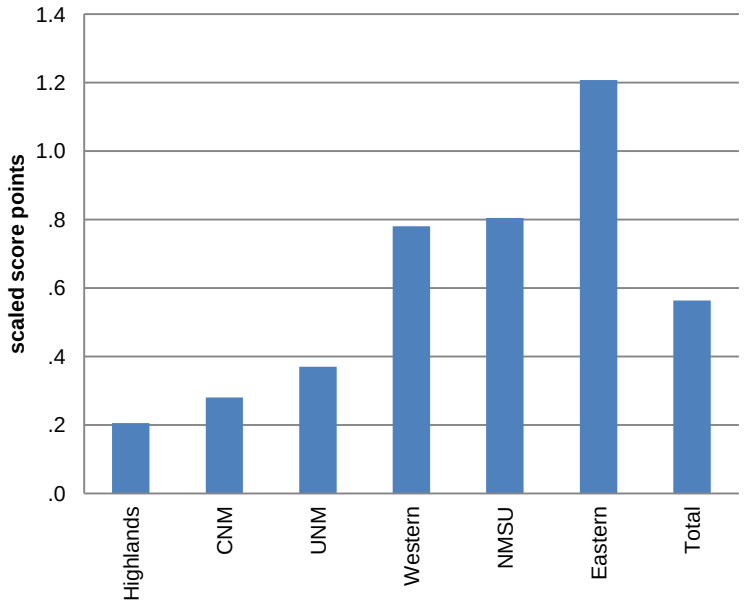
- Program approval currently not based on outcomes data
- Value-added results are one source of data on preparation effectiveness
- Value-added scores range from -0.5 points to 0.4 points



Source: LFC Analysis

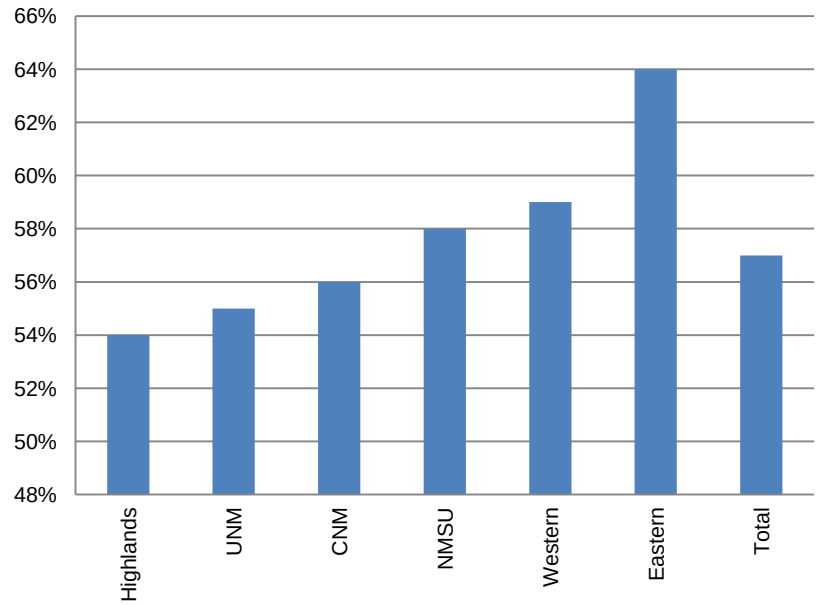
Finding Two: Other Outcome Measures

Chart 16. Average Reading Scaled Score Gains, SY11 to SY12



Source: LFC analysis

Chart 18. Students with at Least One Year of Reading Growth SY11 to SY12

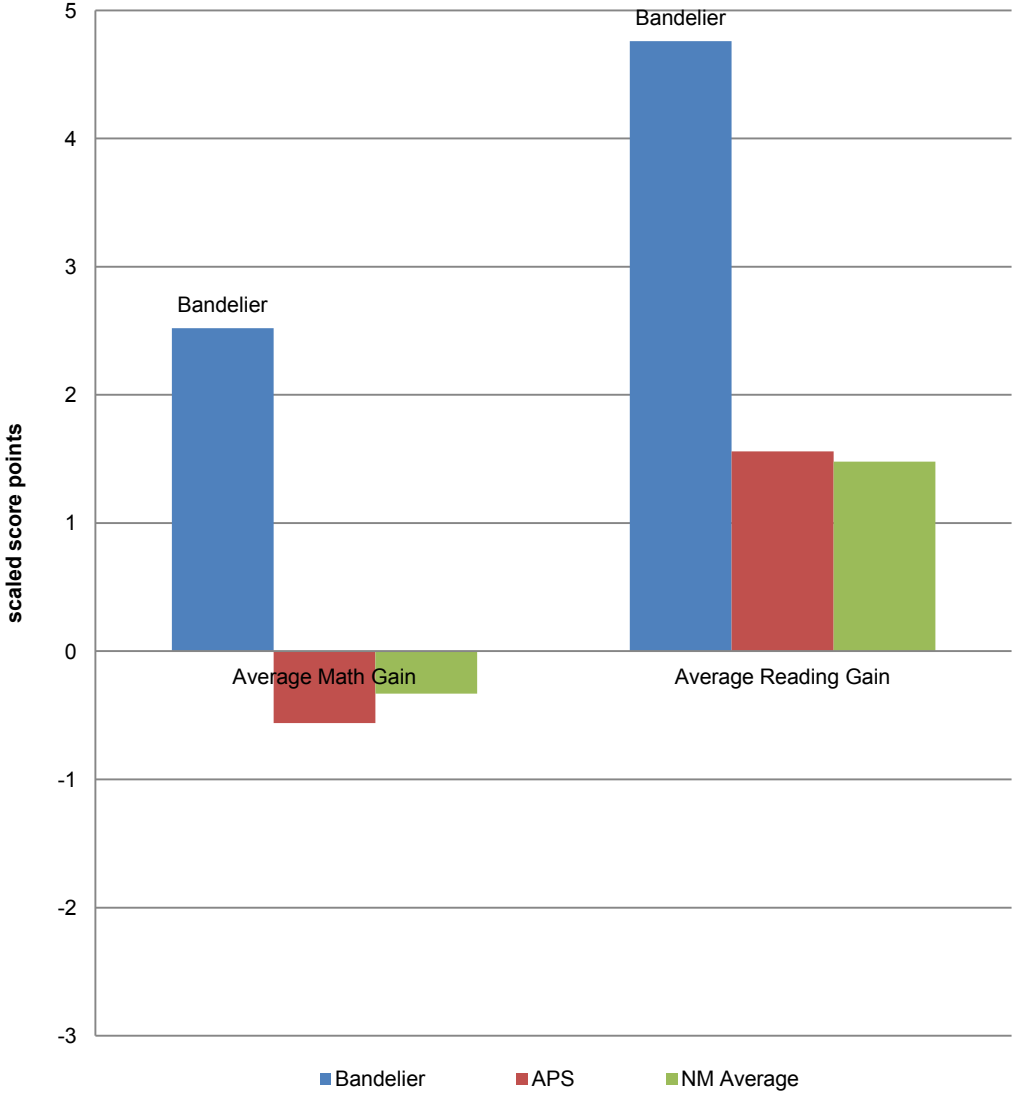


Source: LFC Analysis

Finding Two: Improving Program Quality

- Teachers least prepared to meet the needs of students with disabilities, teach English language learners (ELL), and effectively use student data
- Syllabi identified opportunities for improvement as well as promising practices
- New reading assessment leading to course revisions
- Importance of high quality field experience- Bandelier model

Chart 21. Bandelier Fifth-Grade Scaled Score Increases, SY12



Source: LFC Analysis

Finding Three: Administrator Preparation

Table 14. New Mexico Educational Administrator Assessment Pass Rates, 2008 - 2010

University	First-time Pass Rate	Average Scaled Score
Eastern	100%	NR
Highlands	NR	NR
NMSU	NR	NR
UNM	100%	271
Western	100%	264.5
Statewide	**100%**	**262.3**

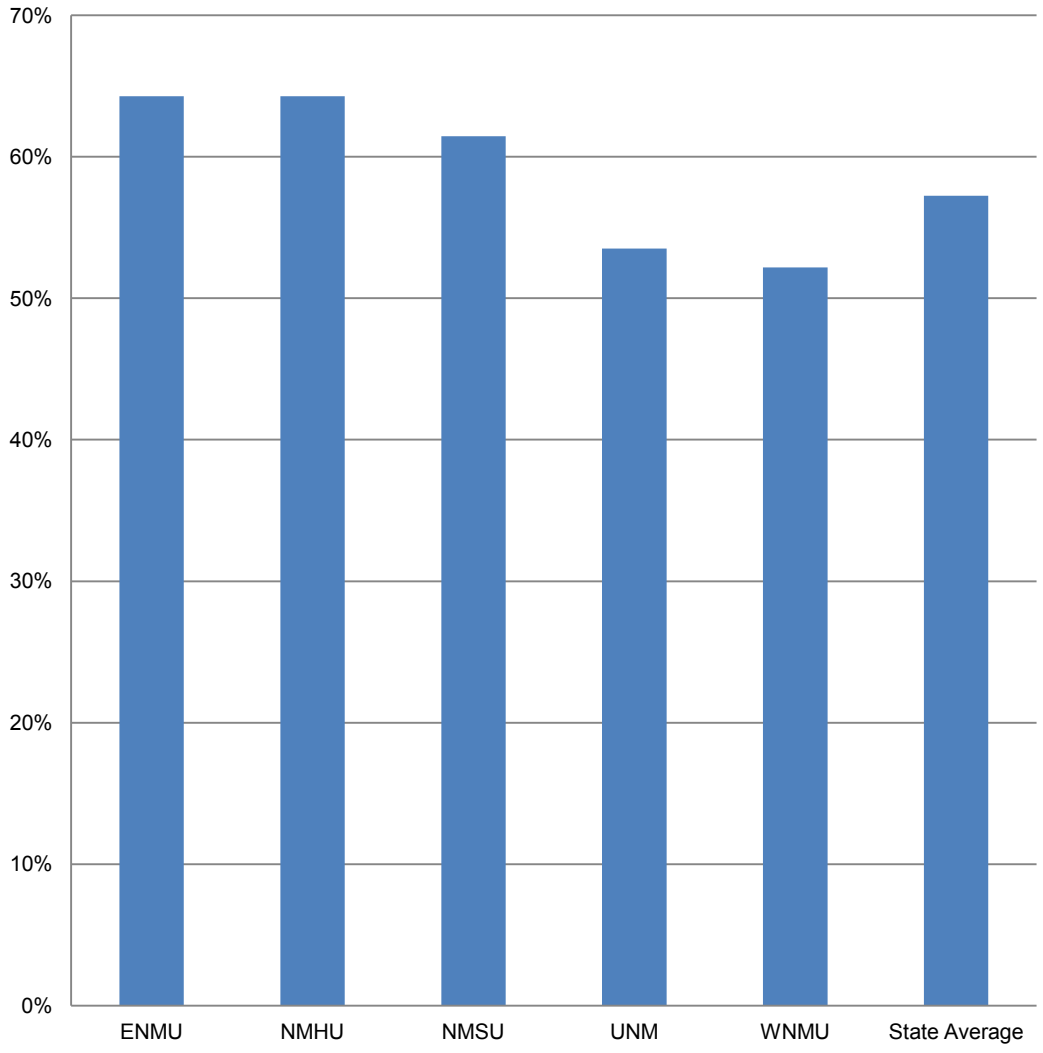
Source: 2011 Title II Report

- Similar to teacher prep: low entrance standards, low exit standards, and opportunities for improving field experience
- Regulatory barriers to the supply of school leaders
- Promising practice: UNM-APS partnership

Finding Four: Improving EARS

- Most work from colleges of education, not PED
- Overlap with Title II
- Opportunity for reporting outcomes data, i.e. persistence rates

Chart 26. Average Three-Year Persistence Rate of Teachers Licensed in 2008 and 2009



Source: LFC Analysis

Key Recommendations: The Legislature

- Couple increases in beginning teacher licensure standards with level I starting teacher salaries beginning in FY16.

- Revise statute to substitute the federal Title II report for the educator accountability reporting system, and include student outcome and teacher retention data by college.

Key Recommendations: The PED

- Phase-in increases to the NMTA licensing cut scores, beginning in FY16.

- With the colleges of education, the LFC, and the LESC, develop a methodology for calculating average value-added scores by institution, calculate this value-added score annually, and identify performance benchmarks for each college of education.

- Consider student outcome data, educator retention data, and school grades in the program approval and renewal process.

Key Recommendations: Colleges of Education

- Raise admissions requirements, including the minimum NMTA basic skills assessment scores.

- Improve and expand research-based teacher and administrator clinical experiences.

Key Recommendations: Higher Education Department

- Incorporate teacher preparation program outcome data and employment retention rates in the higher education performance-based funding formula.

QUESTIONS OR COMMENTS?

LFC Program Evaluation: Promoting Effective Teaching in New Mexico

Presentation to the
Legislative Education Study Committee
July 11, 2013

Overview

- Quality teaching is the most influential school factor affecting academic success.
- Three-tiered system created in 2003
- Previous LFC work confirmed the system met goals to decrease shortages, reduce unqualified teachers, and improve teacher pay.
- Student performance, however, has not improved with taxpayer investments in teacher pay.
- System framework is worth keeping and improving

System Overview

- Progressive career ladder
- Teaching competencies
- Multiple evaluations – uniform standard
- Local Evaluations
  - Annual, focused on competencies, pass/fail
  - 3rd year summative
- State Evaluation
  - Professional Development Dossier
  - Used to screen for licensure advancement and associated pay increases

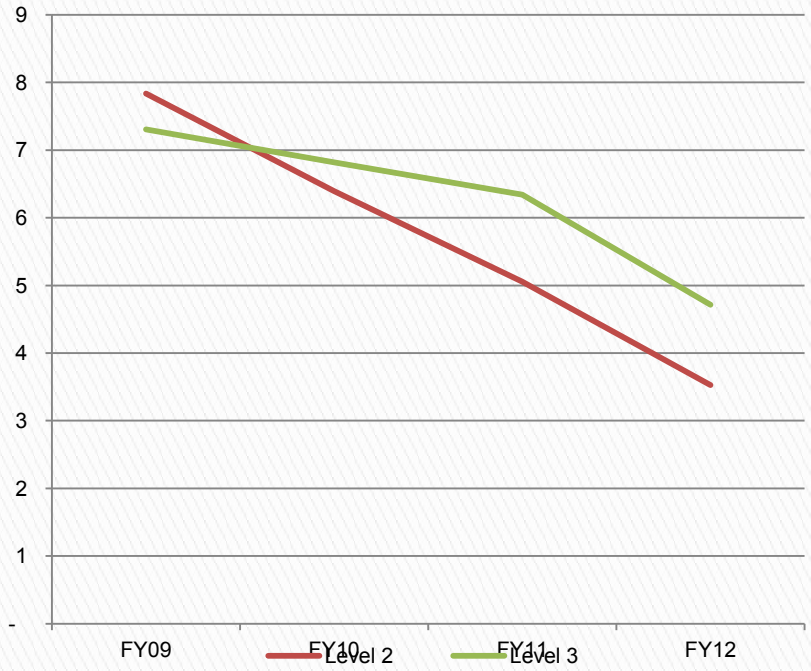
System Status

Table 3. SY09-SY12 Number of Teachers Ascending Licensure Levels

	SY09	SY10	SY11	SY12	Total
From level I to level II	904	1,278	786	909	3,877
From level II to level III	497	637	384	462	1,980
Total	1,401	1,915	1,170	1,371	5,587

Source: LFC Analysis of PED Data

Chart 1. Grandfathered Teachers by Licensure Level
(in thousands)



Source: LFC Analysis of PED Data

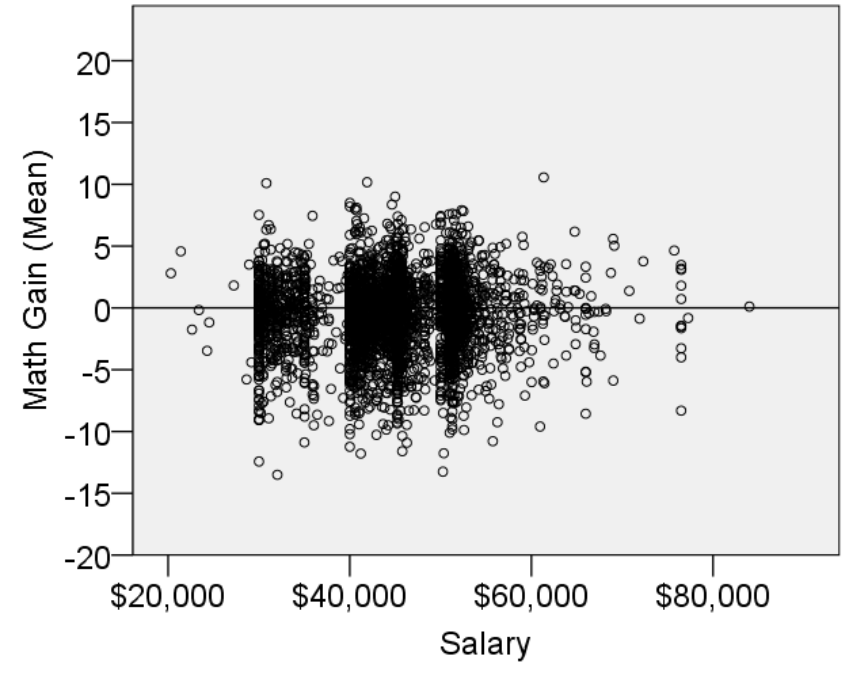
Since 2009, nearly 6,000 teachers advanced to new licensure levels , receiving $59 million in mandatory salary increases.

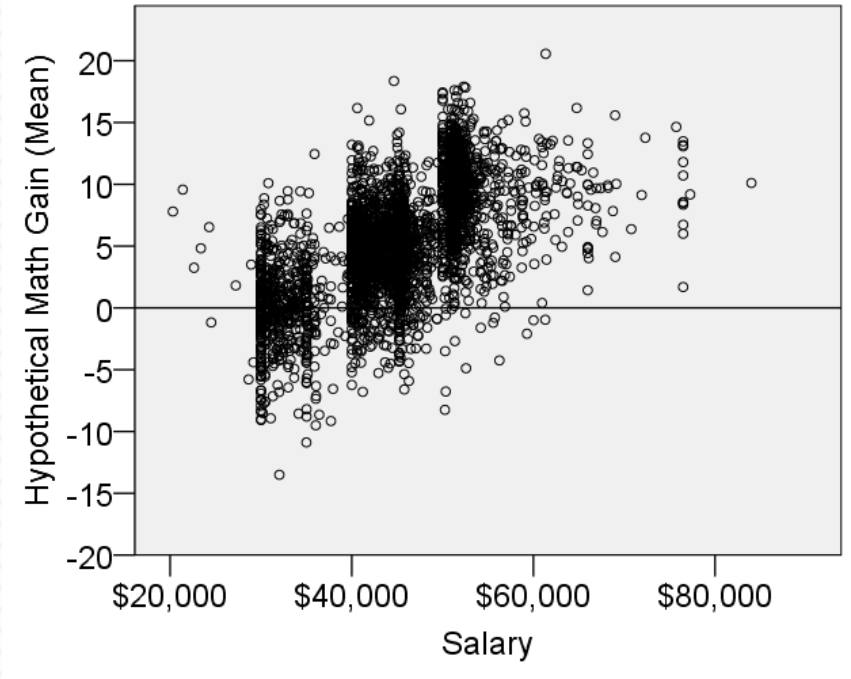
Teachers advancing through the system increasingly make up a larger proportion of classroom teachers than those grandfathered into their licensure level.

Pay Not Aligned With Outcomes

- The state has not established expectations for student achievement across level I, II, and III teachers as part of evaluation systems.
  - *Student achievement is not a robust element of the current three-tiered system.*
  - *Teachers across licensure levels produce similar student achievement results, despite large differences in pay and cost to taxpayers.*
  - *Teachers in higher tiers generally produce better outcomes for students, but these differences are small and can often be accounted for by other factors (fewer low-income and ELL students).*

Performance by Pay (Math)



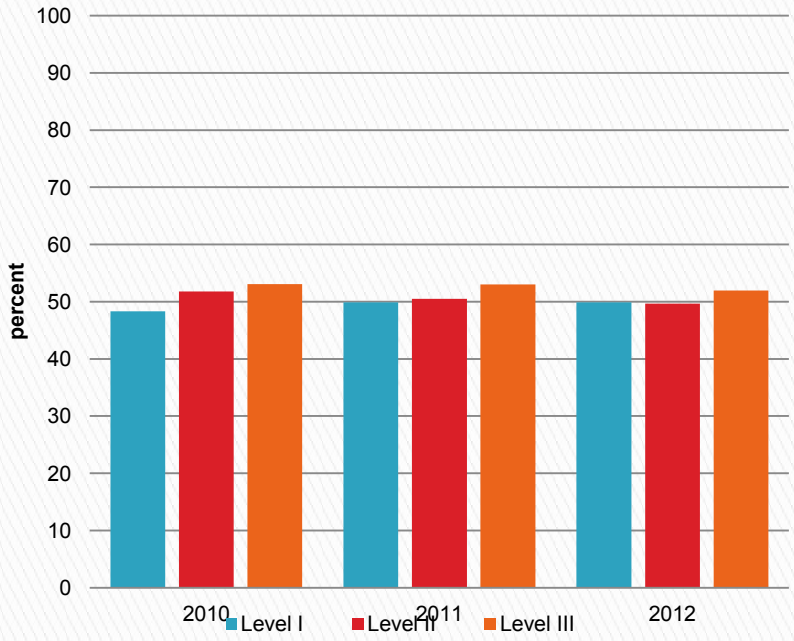


Actual Performance

Hypothetical

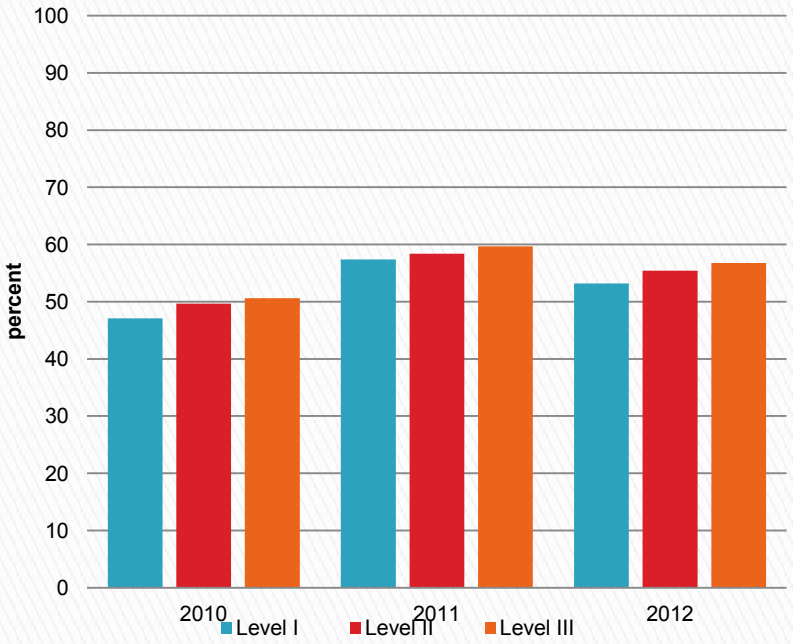
Academic Growth

Chart 4. Percent of Math Students Obtaining One Year of Growth Per Tier (as measured by SBA gain scores)



Source: LFC Analysis of PED data

Chart 5. Percent of Reading Students Obtaining One Year of Growth Per Tier (as measured by SBA gain scores)



Source: LFC Analysis of PED data

Math

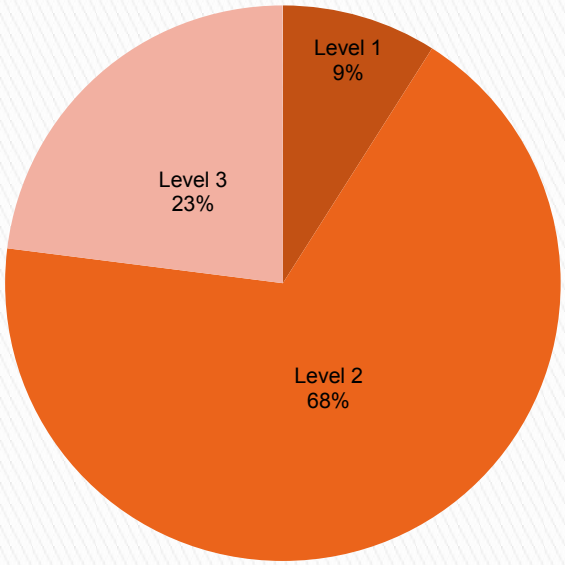
Reading

Improved Evaluations Needed

- The local evaluation system does not differentiate between high and low-performing teachers or focus on student achievement.
- The professional development dossier does not effectively screen teachers for advancement, resulting in ineffective teachers receiving large pay increases.

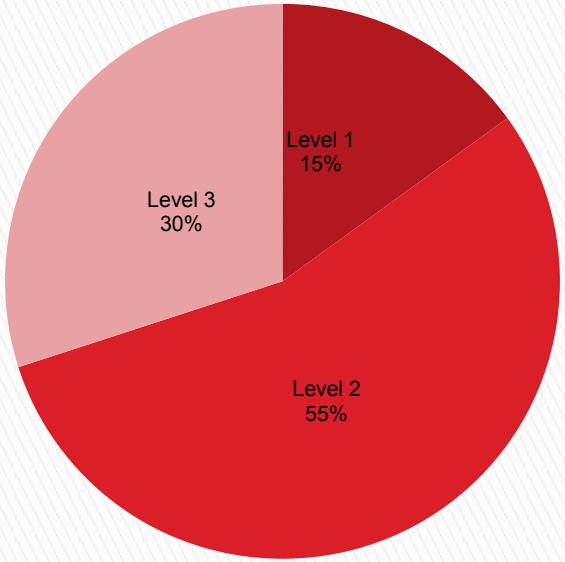
Distribution of High/Low Performing Teachers

Chart 12. FY12 License Levels as a Proportion of High Performing* Reading Teachers



Source: LFC Analysis of PED Data
*High-performing teachers are those ranked in the top 16th percentile of all teachers

Chart 13. FY12 License Levels as a Proportion of Low Performing* Reading Teachers



Source: LFC Analysis of PED Data
*Low-performing teachers are those ranked in the bottom 16th percentile of all teachers

High

Low

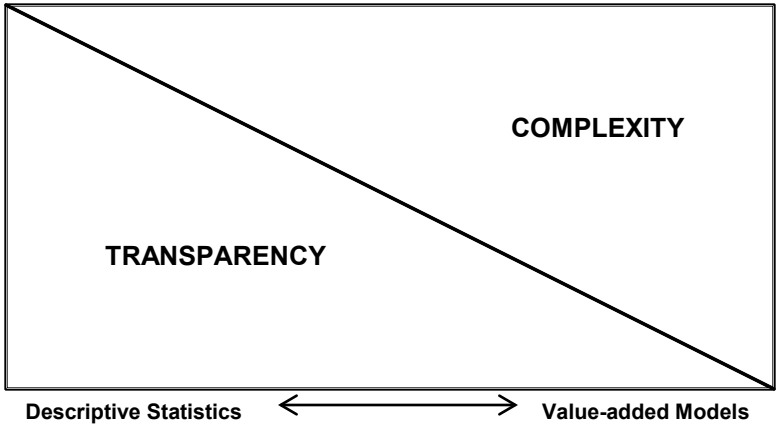
Keep System – Just Update

- The three-tiered system offers a framework to align resource allocation to performance, but student achievement must be better incorporated into the process.
  - *The three-tiered system successfully retained teachers in New Mexico schools.*
  - *PED now captures student achievement data that could be incorporated into the PDD to make it more robust.*
  - *Opportunity exists to incorporate demonstrated effectiveness into passage between tiers and renewal and better align funding and results.*

Value-Added Models

- When used appropriately, value-added models can identify effective teachers and drive student achievement
  - Can identify very high and low ends of performance
  - Ranking system against avg. performance
  - Volatility warrants caution

Figure 1. Understandability of Statistical Models



Model Choice Importance

- Some value-added models adversely affect educators teaching certain populations of students.
- Some VAMs attempt to control for demographic factors and may use multiple years of scores on a handful of different assessments, while others do not.
- In 2012, Pearson Education, Inc. published a study comparing five different VAM teacher evaluation approaches and concluded VAM results are not definitive, and will depend on the model used.

VAM – Examples

Mr. Wilson – Gifted Education Teacher

Mr. Wilson teaches at a large, urban district. In 2011, 97 percent of Mr. Wilson's students were classified as gifted. Twenty percent of his students qualified FRL and 42 percent were Hispanic. As a teacher of mostly gifted students, Mr. Wilson's VAM rating would vary depending on whether student demographic factors were included in the model.

How Different VAMs Affect Mr. Wilson's Status

	Test Score Only Model (no student demographic factors)		Student Demographic Model (includes all available student demographic factors)	
	Math	Reading	Math	Reading
1 year of data	Exceptionally Effective	Highly Effective	Ineffective	Ineffective
2 years of data	Highly Effective	Meets Expectations	Needs Improvement	Needs Improvement

Because Mr. Wilson specializes in teaching a high-performing group of students and improves their student achievement, models that do not incorporate student demographics reflect his effectiveness. Models that do incorporate student demographics penalize Mr. Wilson because his students are gifted and a relatively lower proportion of them are in poverty.

VAM – Examples

Ms. Campos – Teacher of At-risk Students

Ms. Campos teaches 3rd grade in small, rural school district. Over 90 percent of her students are FRL and are classified as English language learners. Additionally, all of her students participate in special education and 100 percent of her students are Native American. As a teacher of this highly at-risk group of students, Ms. Campos' value-added rating in a VAM depends heavily on which model is applied.

How Different VAMs Affect Ms. Campos' Status

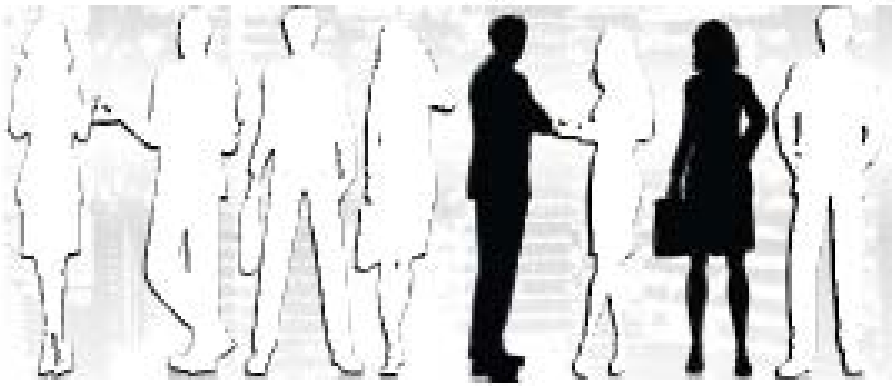
	Test Score Only Model (no student demographic factors)		Student Demographic Model (includes all available student demographic factors)	
	Math	Reading	Math	Reading
1 year of data	Needs Improvement	Needs Improvement	Highly Effective	Highly Effective
2 years of data	Needs Improvement	Meets Expectations	Needs Improvement	Meets Expectations

Ms. Campos moves between one of the lowest performance categories to the highest. Using a model with no student demographic factors could discourage effective teachers from accepting positions in low-income schools.

VAMs Not Available for All

Figure 2. Teachers in New Mexico

Figure 3. Teachers in New Mexico for whom we can compute VAM scores



Using VAMs in 3-Tiered System

- The use of VAMs can be responsibly integrated into the three-tiered system to identify teachers for advancement and bonus pay.
  - *VAMs can do a good job of identifying highly effective and highly ineffective teachers for rewards and interventions.* (Advancement)
  - *VAMs should not be used in annual local evaluations because of their limitations and complexity. (High Stakes)*

Recommendations

- Strengthen the current three-tiered system, including licensure advancement process to ensure teachers demonstrate performance commensurate with higher pay. Ensure continued teacher effectiveness by strengthening licensure renewal.
- Establish expectations for the use of student performance data that is transparent, fair and useful for improving professional practices and improved results for students.
- Align state funding with highly effective teaching through the funding formula.
- See Handout

Potential Statutory Changes from LFC Staff Report on Effective Teaching

Framework for Assessing and Rewarding Teacher Effectiveness: Summary

- Strengthen the current three-tiered system, including licensure advancement process to ensure teachers demonstrate performance commensurate with higher pay. Ensure continued teacher effectiveness by strengthening licensure renewal.
- Establish expectations for the use of student performance data that is transparent, fair and useful for improving professional practices and improved results for students.
- Align state funding with highly effective teaching through the funding formula (separate legislation).

Annual Evaluation

Strengthen statutory requirements for a highly objective uniform standard of evaluation (HOUSE) for teachers by requiring the following.

- Establish Updated Basic Competency and Effectiveness Indicators for Teachers.
- Require Professional Development Plan by 40th day establishing the current year's performance goals, including measurable objectives for student performance. The goals should be based on Basic Competency and Effectiveness Indicators, the previous year's annual evaluation, and previous year's students' performance.
- Performance Evaluation: Annual evaluation should be based on whether the teacher met or exceeded expectations on Basic Competency and Effectiveness Indicators, made satisfactory progress on professional development goals, and received satisfactory ratings from students and parents. Require classroom observations from principals.
- Performance Improvement Plan: Establish a structure to provide assistance to teachers not meeting expectations.
- Performance Remedy: Clarify that public school remedy for non-performance includes non-renewal of contract, or other action (suspension or termination) in accordance with other existing due process laws.
- Require that only teachers meeting or exceeding expectations on annual performance evaluations may receive state or district funded salary increases the subsequent year.
- Training: Require principals to receive training at least once every two years to improve evaluation skills.
- Local Schools: Require local schools to implement policies and procedures to implement this section and authorize PED to approve supplemental options and

measures for a local system of data collection for the annual teacher performance evaluation, including the use of peer observations, to improve teaching practice.

Summative Effectiveness Evaluation

- After three years of classroom teaching require an effectiveness evaluation to be conducted no later than the 40th day the following school year and include 3-year summaries of progress meeting Basic Competency and Effectiveness Indicators, achieving professional development goals, and improving student achievement. Improving student achievement component should count for no less than 50 percent an overall rating.
- The summative effectiveness evaluation includes a cumulative assessment of a teachers' effectiveness at improving student achievement over time, as measured by PED expected student performance growth targets on approved assessments. Performance expectations should be aligned with the three-tiered licensure levels, and subject and grade level standards.
- Public schools may award teachers with successful effectiveness evaluations multi-year contracts not to exceed the equivalent term of a contract of the district's superintendent. Public schools may use the results of the effectiveness evaluation to make employment decisions, in accordance with other provisions of law.

Three-Tiered Licensure System

License Term

- Level I – 5 years, however a teacher must submit for advancement after three years.
- Level II and Level III – 8 years, however a teacher must submit for renewal after six years.

Advancement

- Level I to Level II
    - Require three years of classroom teaching at Level I before advancement.
    - Require one year of mentor program.
    - Require three years of satisfactory annual evaluations.
    - Superintendent approval of advancement and verification of submittal information.
    - Meet performance expectations as demonstrated through Effective Teaching Portfolio –OR- Performance Ranking.
        - Require meet or exceed score on Effective Teaching Portfolio
        - Performance score ranking using PED-approved Value-Added Model in the top 50 percent of Level II teachers statewide according to standards-based assessment student achievement data.
- Level II to Level III
    - Require three years of classroom teaching at Level II before advancement.

- Require satisfactory effectiveness evaluation for most recent three year period.
- Superintendent approval of advancement and verification of submittal information.
- Meet performance expectations as demonstrated through Effective Teaching Portfolio –OR- Performance Ranking.
  - Require meet or exceed score on professional development dossier
  - Performance score ranking using PED-approved Value-Added Model in the top 50 percent of Level III teachers statewide according to standards-based assessment student achievement data.

Renewal

- Level II
  - Satisfactory score on effectiveness evaluation for most recent three year period.
  - Satisfactory score on student achievement portion of Effective Teaching Portfolio –OR- VAM in top 50 percent of Level II teacher rankings statewide.
  - A teacher failing to meet renewal requirements within license term may apply for a provisional Level II license and demonstrate satisfactory performance within two years.
- Level III
  - Satisfactory score on effectiveness evaluation for most recent three year period.
  - Satisfactory score on student achievement portion of Effective Teaching Portfolio –OR- VAM in top 50 percent of Level III teacher rankings statewide.
  - A teacher failing to meet renewal requirements within license term may apply for a provisional Level III license and demonstrate satisfactory performance within two years. Level III minimum salary requirements are waived for provisional Level III teachers and are subject to local pay policy.

Effective Teaching Portfolio (replace Professional Development Dossier)

- Require submission of an Effective Teaching Portfolio (ETP) as part of the licensure advancement application. The ETP should include evidence of effective teaching practice in three areas: instruction, professional development and student learning. Evidence of student learning should constitute at least 50 percent of the overall ETP score.
  - Instruction includes evidence of instructional plans, assessment techniques, use of data to inform practice, adaptation of teaching for diverse learners, classroom management, and implementation of state content standards.

    - Professional development includes evidence of meeting professional development goals, collaborating with other educators, parent involvement, or research and publication.
    - Student learning includes evidence of improved student achievement on PED approved assessments using at least three years of data.
- The ETP should be evaluated by two external reviewers, one of whom shall hold the same grade level licensure and subject area endorsement as the candidate.

Performance ranking through value-added model scores

- Require PED to annually rank the performance of licensed teachers providing instruction in tested grades and subjects and provide results to public schools and the individual teachers. The results should benchmark performance relative to teachers statewide, within the district, the school and license level by grade taught and overall. The results should be part of a teacher's personnel file, confidential and only available for review by authorized personnel.
- Rankings should be determined using at least three years of standards-based student performance data and a composite score of a simple and a complex value-added model.

Transparency (additional possible measures)

- Require PED to establish a Teacher Effectiveness Advisory Council, and charge it will recommending Basic Competency and Effectiveness Indicators for use in the annual evaluation, expected student performance growth targets and assessments that can be used for the Effectiveness Evaluation, and Value-Added Models for use in lieu of an Effective Teacher Portfolio for licensure advancement.
- Provide PED specific rule making authority for the evaluation and three-tiered sections, require input from the Teacher Effectiveness Advisory Council before publishing proposed rules, and provide a Sunset date for these rules which would require legislative reauthorization of PED rules (this is a new concept in NM).
- Require PED to annually publish student performance growth targets and approved assessments and calculations, including a technical manual, used to compute VAM composite scores.

Funding Formula Alignment and Highly Effective Teacher Awards

- Align funding formula to Three-tiered salary increments over next two years – separate bill.

LFC Program Evaluation Progress Report

Report Issued:
July 2012 • #12-05

Progress Report:
May 2013

KEY FINDINGS

1) Early education improves performance, but lack of coordinated resources and inconsistent quality limits success.
2) Early literacy initiatives, such as mandatory retention policies and Reading First, have produced mixed results
3) State, district, and school-level management policies can help schools to marginally beat the odds.

Third-Grade Achievement Gap Related to Poverty, 2011

	Free Lunch	Non-Free Lunch	Diff.
Hisp.	44%	68%	24%
Caucas.	61%	80%	19%
Nat. Amer.	33%	55%	22%

Source: LFC Analysis

Of 25 thousand third-graders in New Mexico in 2011, only 81 attended PreK and two years of K-3 Plus.

Developing Early Literacy in New Mexico (PED)

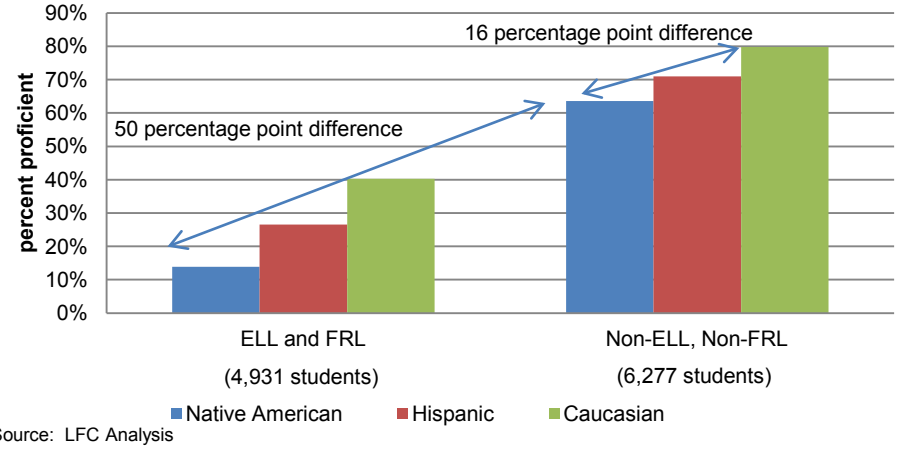
Determined reading proficiency rates over time and relationships to student demographics, evaluated spending patterns and practices the state and districts use to finance early literacy, and analyzed best practices for accelerating student achievement in literacy.

Background: Early reading proficiency is well-established as a strong predictor of high school graduation rates as well as future earning potential. In spite of slight improvements to scaled scores on New Mexico's standards-based assessments (SBA), third-grade reading proficiency rates continue to lag behind desired levels. In response, New Mexico has invested heavily to improve early literacy, including full-day kindergarten, prekindergarten (PreK), and an extended school- year program, Kindergarten-Three Plus (K-3 Plus). The Legislature has quadrupled funding for PreK since FY06 and doubled funding for K-3 Plus between FY12 and FY13.

Key Recommendations:

- PED should annually report its process for determining SBA cut scores, evaluate the impact of bilingual models on New Mexico's English Language Learners, and raise attendance criteria in school grading to encourage schools to improve performance rates.
- PED should track enrollment in PreK and K-3 Plus to increase the number of students who receive benefits and increase oversight and accountability of K-3 Plus to improve consistency and quality of implementation.
- PED should adopt short-cycle assessments in grades K-3 that align to the Common Core standards
- PED and CYFD should consider alternative PreK assessments based on: cost effectiveness, time required for administration, and alignment to the Common Core.

FRL and ELL Third-Grade Reading Achievement Gap, FY11



Source: LFC Analysis

LFC Program Evaluation Progress Report

Ten percent of New Mexico's third-grade class in 2011 had been retained between kindergarten and third-grade, but only 29 percent of those students were proficient as third-graders.

Retained Third-Grade SBA Reading Score Change from SY10 to SY11

SBA Reading Point Change from SY10 to SY11	% of Retained Students
≤ 0	14%
1-4	18%
5-8	18%
9-11	15%
12-14	11%
15-18	11%
>18	12%

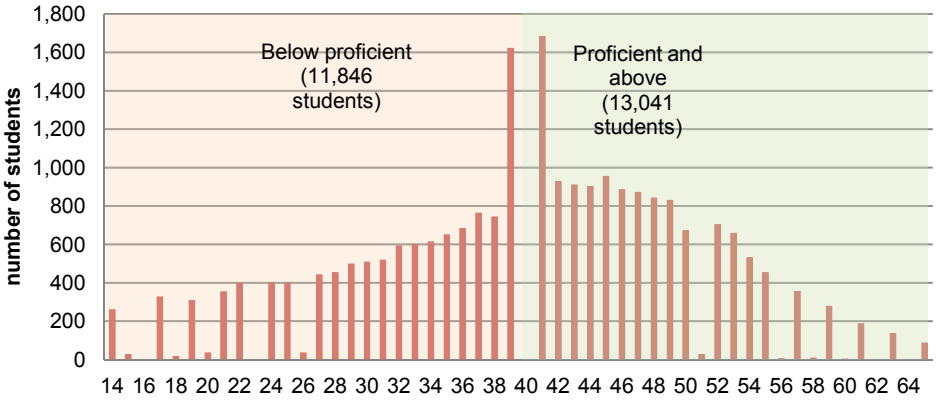
Source: LFC Analysis

In 2011, 2,446 third-graders, or 10 percent, were within 2 points of scoring proficient on the standards-based assessment.

Agency-reported Progress to Date:

- For FY14, $16 million was appropriated for K-3 Plus. As of June 2013, $15.3 million was allocated to districts for July-August 2013 K-3 Plus programs. The remaining FY14 funds will be allocated to districts for June 2014. In summer 2013, the projected enrollment is 16 thousand students, an increase of 9,000 students. In addition, 65 more schools and 12 more districts are participating than in 2012.
- For FY14, $15 million was appropriated for PreK. PED allocated $13.7 million to districts, charter schools, and RECs to provide PreK services for 4,230 children in 98 school sites. In FY13, $10 million was appropriated to provide PreK services for 2,850 children at 67 school sites.
- As part of New Mexico's $25 million federal Race to the Top early learning grant in 2012, PED will develop unique identifiers for children in Head Start and childcare programs.
- PED's bilingual division will disaggregate ELL program data in 2013.
- A committee meets twice per year to advise PED on K-3 Plus.
- Beginning in FY13, PED adopted a K-3 formative assessment, DIBELS Next, for screening and progress monitoring of early literacy skills.

Third Grade Reading SBA Scaled Score Distribution, FY12
(24,887 students)



Source: PED
Note: Excludes 624 students scoring 0-13 and 66-80.

Outstanding Issues:

- As part of the FY14 budget submissions, PED required districts to report data on student attendance and identify strategies for improvement; the department is developing next steps for expanding those strategies statewide.
- PED and CYFD are yet to consider an alternative PreK assessment to indicate readiness for kindergarten.
- The PED is yet to report its process for determining SBA cut scores.

Report
to
The LEGISLATIVE FINANCE COMMITTEE

Public Education Department
Teacher and Administrator Preparation in New Mexico
December 5, 2012

Report #12-13

LEGISLATIVE FINANCE COMMITTEE

Senator John Arthur Smith, Chairman
Representative Luciano "Lucky" Varela, Vice-Chairman
Senator Sue Wilson Beffort
Senator Pete Campos
Senator Carlos R. Cisneros
Representative William "Bill" J. Gray
Senator Stuart Ingle
Representative Rhonda S. King
Representative Larry A. Larrañaga
Senator Carroll H. Leavell
Senator Mary Kay Papen
Representative Henry "Kiki" Saavedra
Representative Nick L. Salazar
Representative Edward C. Sandoval
Senator John Sapien
Representative Don L. Tripp
Representative James P. White

DIRECTOR

David Abbey

DEPUTY DIRECTOR FOR PROGRAM EVALUATION

Charles Sallee

PROGRAM EVALUATION TEAM

Jeff Canney, CGFM
Jon Courtney, Ph.D.
Valerie Crespin-Trujillo
Jack Evans
Brenda Fresquez, CICA
Pamela Galbraith
Maria Griego
Rachel Mercer-Smith
Matthew Pahl
Michael Weinberg, Ed.D.

State of New Mexico
LEGISLATIVE FINANCE COMMITTEE
325 Don Gaspar, Suite 101 • Santa Fe, NM 87501
Phone: (505) 986-4550 • Fax (505) 986-4545

David Abbey
Director

December 5, 2012

Ms. Hanna Skandera, Secretary-Designate
Public Education Department
Jerry Apodaca Education Building
300 Don Gaspar Avenue
Santa Fe, NM 87501

Dear Ms. Skandera:

On behalf of the Legislative Finance Committee (Committee), I am pleased to transmit the evaluation, *Teacher and Administrator Preparation in New Mexico.* The program evaluation team followed-up on the 2006 LFC evaluation of colleges of education, reviewed the status of New Mexico's educator accountability reporting system, and analyzed relationships between teacher and administrator education programs and student performance. The report will be presented to the Committee on December 5, 2012. Exit conferences were conducted with the Public Education Department on November 20, 2012 to discuss the contents of the report. The Committee would like a plan to address the recommendations within this report within 30 days from the date of the hearing.

I believe this report addresses issues the Committee asked us to review and hope New Mexico's public education system benefits from our efforts. We very much appreciate the cooperation and assistance we received from your staff.

Sincerely,

David Abbey, Director

Cc: Senator John Arthur Smith, Chairman, Legislative Finance Committee
Representative Luciano "Lucky" Varela, Vice-Chairman, Legislative Finance Committee
Representative Henry "Kiki" Saavedra, Legislative Finance Committee
Representative Rick Miera, Chairman, Legislative Education Study Committee
Ms. Frances Maestas, Director, Legislative Education Study Committee
Dr. Tom Clifford, Secretary, Department of Finance and Administration

Table of Contents

Page No.

Number of Licensed Teachers by College of Education*

	Licensed Teachers	Percent of Total
CNM	420	1%
Eastern	1,792	4%
Highlands	1,333	3%
NMSU	2,859	6%
UNM	5,368	11%
Western	815	2%
Other**	13,077	28%
Unknown	21,758	46%
Total	**47,422**	

*Not all licensed teachers are actively teaching.
**Includes teachers prepared by private, out-of-state, and other in-state institutions.
Source: LFC Analysis

The state has invested $59 million in mandatory salary increases through the three-tiered system since 2009.

Students Proficient or Above on the SBA, SY12

	3rd Grade	8th Grade	11th Grade
Reading	52%	54%	45%
Math	53%	42%	39%

Source: PED

The Public Education Department has not established expectations for how well teachers should improve student performance.

Teachers and principals are the most important school-based factors affecting student learning, and New Mexico's six largest colleges of education prepare half of the state's licensed educators. Colleges of education account for 11 percent of the state's student credit hours, generating $64 million in higher education funding formula revenue. Additionally, in FY12, districts and charter schools budgeted $1.2 billion for teacher salaries and benefits, making up 50 percent of K-12 formula funding and 22 percent of total general fund appropriations. Given that more than half of K-12 students in New Mexico perform below grade level, it is vital that the state's colleges of education prepare high quality teachers and administrators.

In 2006, the Legislative Finance Committee (LFC) evaluated five teacher preparation programs in New Mexico, finding revenues exceeded expenditures at each program, low percentages of full-time faculty, lower requirements for field work than what is considered best practice, and low requirements for passing scores on the New Mexico Teacher Assessments. While the colleges of education implemented recommendations to develop and improve the educator accountability system, minimal programmatic changes occurred and student achievement has remained disappointingly low. This evaluation assesses the progress made to implement previous recommendations, including the educator accountability system, and analyzes the relationship between teacher and administrator programs and student performance.

While slight differences exist between programs, the overall performance of teachers lags behind what is necessary to help students make "catch-up" growth. These student outcomes are partially related to low entry and licensure standards, despite attempts to attract high-quality teachers through the three-tiered licensure system. By more closely overseeing teacher quality on the front end, the Public Education Department (PED) can reduce the burden of dealing with ineffective teachers through evaluations and professional development.

This report highlights the importance of carefully selecting candidates for teacher and administrator preparation programs, raising licensure standards for educators, actively monitoring the performance of preparation programs, and connecting the higher education funding formula to educator quality. Using outcomes data, including K-12 standardized test scores and teacher retention rates, this evaluation identifies effective practices within the state's colleges of education worth replicating statewide. These include coursework changes as well as improvements to fieldwork experiences for both teachers and administrators.

Prior to admission, teacher candidates must demonstrate academic skills generally acquired during middle school. The test's passing score is set far below average.

Several states, including Massachusetts, Pennsylvania, and Tennessee, recently raised cut scores for their teacher competency exams.

Failed Elementary Assessment Attempts Before Passing 2002 – 2012

Number of Failures	Number of Teachers
1 - 5	326
6 - 10	19
11 - 17	3
Total	**348**

Source: LFC Analysis

Teachers who scored 260 on the math content assessment are predicted to add an average of 1.4 points to their students' SBA scaled scores compared with teachers who earned a minimum passing score of 240.

Highlands remained on PED's list of approved teacher preparation programs in spite of losing accreditation between 2007 and 2012.

KEY FINDINGS

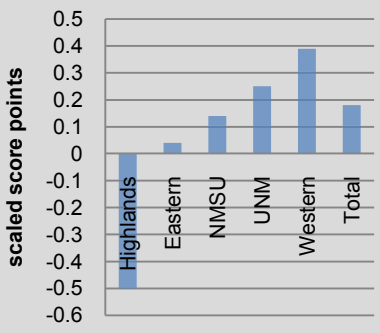
Low teacher admission requirements and licensure standards perpetuate low student performance Despite investments in the state's three-tiered licensure system, colleges of education continue to attract and admit academically average candidates. While the state's colleges of education do not require minimum ACT scores for admissions, the average scores of teacher candidates have not increased since the 2006 LFC evaluation.

New Mexico's teacher competency exams provide little information about program quality as virtually all teachers pass. Since 2008, every institution's passage rates exceed 90 percent on the basic skills test, elementary competency test, and secondary competency test. Since being set by the State Board of Education in 2000, the passing score for all NMTA assessments remains at 240, one standard deviation below the average score of 260.

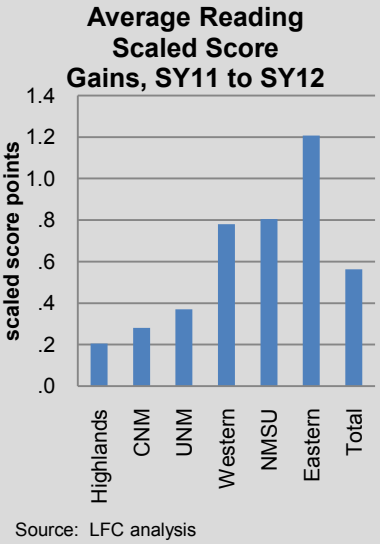
Teachers who fail an NMTA at least one time perform lower than those who pass on their first attempt. As noted in previous LFC evaluations, one way of measuring a teacher's effectiveness is calculating the difference between how well that teacher's students performed compared with expected performance. Using these value added scores, teachers who failed the elementary content knowledge assessment at least one time added less value to their students, -0.23 points, than those who passed on their first attempt, 0.3 points. Similarly, teachers who score higher on the basic skills assessment, the elementary content knowledge assessment, and the mathematics content knowledge assessment improve student achievement at higher levels.

Raising cut scores would require higher performance from prospective teachers, although New Mexico's teaching supply can withstand increases to licensure standards. New Mexico's teacher preparation programs currently supply an adequate number of completers to replace educators leaving the profession. School districts reported 1,810 teachers left the workforce between SY11 and SY12, while New Mexico's colleges of education prepared 1,277 teacher candidates during SY10. Given that half of the state's teaching force is prepared in-state, this rate of preparation currently exceeds the need. Additionally, 26 thousand out of 47 thousand licensed teachers, or 56 percent, were not actively teaching during SY12, providing a significant eligible reserve of teachers.

The Public Education Department could better oversee preparation programs to improve teacher quality. The PED does not use student and teacher outcome data to approve and renew educator preparation programs, unlike at least six other states that use value-added data to evaluate effectiveness. Given New Mexico's low proficiency rates, moving large numbers of students to grade-level performance will require significant gains. For example, even making two points of progress per year, it will take a student at least five years to move from the lowest performance level



Source: LFC Analysis



Source: LFC analysis

The 4,000 teachers the LFC surveyed referenced student-teaching and hands-on fieldwork as the courses that most prepared them for success.

to proficient. The PED has not quantified the amount of gains it expects of beginning teachers, exemplary teachers, or preparation programs.

Average value-added scores by college range from -0.5 points to 0.4 points, indicating need for overall improvement to increase student achievement. Four of New Mexico's colleges of education have positive value-added scores, while one's value-added score is negative. Alternatively licensed teachers' value-added scores, 0.4 points, are slightly higher than traditionally licensed teachers' average of 0.3 points. Similarly, looking at student scaled score gains between years also highlights differences between programs.

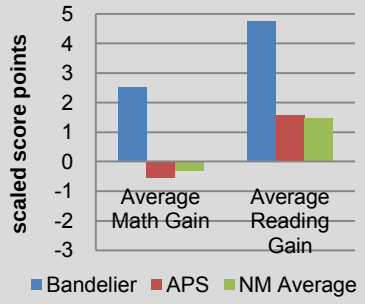
Practitioners and employers agree about recent program completers' areas of weakness, many of which could be better-addressed through coursework. According to LFC survey data, teachers report feeling least prepared to meet the needs of students with disabilities, teach English language learners, and effectively use student data. An LFC review of each college of education's syllabi identified opportunities for improvement as well as promising practices. Western and Eastern, for example, both require courses focused on use of data, while Western and Central New Mexico require all teacher candidates to complete a classroom management course. Many programs are revising reading courses based on a newly implemented licensure exam intended to measure teachers' readiness in the science of reading instruction.

High quality fieldwork produces positive student outcomes. According to an LFC survey of over 200 principals, 80 percent strongly agreed that student teaching is a critical element of teacher preparation, and 86 percent strongly agreed that strategies for effective classroom management, which are often practiced through student-teaching, are critical. However, teacher candidates are not always placed in high-quality professional-development school settings, and placement within clinical school sites often do not persist throughout fieldwork courses. UNM's Bandelier Elementary student-teaching program implements several research-based practices, including extensive collaboration, co-teaching, and selective practicum placement. Though only in its second year, Bandelier shows gains greater than the district average and high rates of teacher placement upon completion.

Increasing entrance standards, exit standards, and programmatic quality will raise administrator quality. Currently, only UNM and Western require recommendations regarding leadership potential. Additionally, selection currently focuses on years of teaching experience, rather than measures of instructional effectiveness described in previous LFC evaluations. Similar to the exams required of teachers, the administrator assessment is not an accurate indicator of preparedness, as between 2008 and 2010, 100 percent of administrator program completers passed.

As measured by school grades, differences in the quality of principal preparation are minimized when student poverty is taken into account. When comparing schools' total grade values and student growth values,

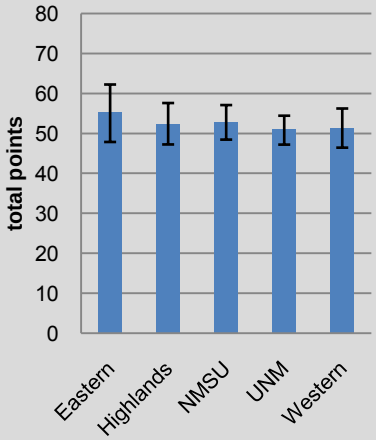
Bandelier Fifth-Grade Scaled Score Increases, SY12



Source: LFC Analysis

Administrator preparation programs are not attracting and selecting candidates with the greatest leadership potential.

Poverty-Adjusted School Grade Totals Among Administrator Preparation Programs



Source: LFC Analysis

Ten of the first 12 completers of the first APS/ UNM leadership cohort, or 83 percent, are employed as administrators, seven times the statewide placement rate of 12 percent.

statistically significant differences appear between programs. For example, Western's principals have lower total school grade scores, 47.2 points, than principals prepared by other in-state programs, 53.9 points. After controlling for school poverty levels, however, school grade differences attributed to administrator programs shrink. Despite the overlap in school performance, practitioners and district administrators perceive school leader preparation programs differently. Based on an LFC survey, principals from NMSU and UNM report the highest levels of preparation, while district administrators most highly rate the preparation of UNM and Eastern graduates.

UNM's principal preparation partnership with APS is a promising clinical practice worth replicating. While New Mexico's colleges of education aligned coursework with the *Interstate School Leaders Licensure Consortium* (ISLLC) leadership standards in 2009, significant differences exist in the quality of the internships the programs require. UNM is partnering with the Albuquerque Public Schools and the New Mexico School Leadership Institute to create a preparation program that includes careful selection of candidates; coursework co-taught by Albuquerque administrators; full-time, semester-long residencies; and follow-up mentoring. Although the program is too new to measure the performance of these leaders' schools, initial placement rates are much higher than the state average.

New Mexico's educator reporting system can be simplified and improved by including outcomes data. While colleges of education have made progress since the 2006 LFC evaluation to develop an educator accountability reporting system (EARS) to provide the state with information about program performance, the report focuses on inputs that overlap with federal reports. Colleges of education consider the duplicate processes redundant and burdensome; the PED does not appear to rely upon EARS to assess how well the state is preparing educators; and the colleges lack access to outcomes data, such as student performance and employment retention rates. PED, however, can calculate employment retention rates and student achievement, which will encourage the colleges of education to focus on producing effective teachers who remain in the profession longer.

KEY RECOMMENDATIONS

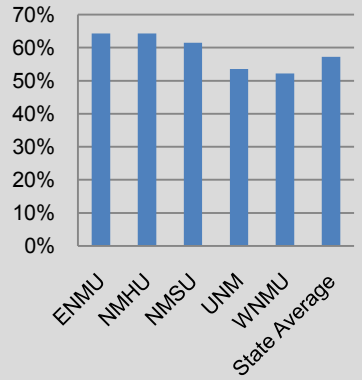
The Legislature should:

Couple increases in beginning teacher licensure standards with level I starting teacher salaries beginning in FY16.

Revise statute to substitute the federal Title II report for the educator accountability reporting system, and include student outcome and teacher retention data by college.

While statute requires PED and colleges of education to collaborate to develop the EARS report, only colleges of education have undertaken this task.

Average Three-Year Persistence Rate of Teachers Licensed in 2008 and 2009



Source: LFC

At Eastern, all elementary candidates complete coursework to receive regular and special education licensure.

The Public Education Department should:

Phase-in increases to the NMTA licensing cut scores, beginning in FY16.

With the colleges of education, the LFC, and the LESC, develop a methodology for calculating average value-added scores by institution, calculate this value-added score annually, and identify performance benchmarks for each college of education.

Consider student outcome data, educator retention data, and school grades in the program approval and renewal process.

Colleges of education should:

Raise admissions requirements, including the minimum NMTA basic skills assessment scores.

Improve and expand research-based teacher and administrator clinical experiences.

The Higher Education Department should:

Incorporate teacher preparation program outcome data and employment retention rates in the higher education performance-based funding formula.

In FY12, districts and charters budgeted $1.2 billion for teacher salaries and benefits, making up 50 percent of K-12 program costs and 22 percent of total general fund appropriations. Statewide, colleges of education account for 11 percent of student credit hours, generating $64 million in formula revenue. This evaluation focused on New Mexico's six largest colleges of education which prepare half of the state's licensed teachers and administrators.

Table 1. Number of Licensed Educators by College of Education

University	Initial Licensure Completers 2010	Percent of Total	Administrative Licensure Completers 2010	Percent of Total
CNM	102	9%	NA	NA
Eastern	100	9%	8	6%
Highlands	107	9%	35	27%
NMSU	349	30%	44	34%
UNM	427	37%	23	18%
Western	70	6%	20	15%
Total	**1,155**		**130**	

Source: LFC Analysis

Workforce trends make teacher and administrator preparation particularly critical. Nationally, the teaching population is slowly aging, and Ingersoll and Merrill (2010) predict teacher retirement will peak between 2011 and 2012. LFC analysis of Education Retirement Board data indicate 2,548 licensed New Mexico teachers, or 9 percent, retired in 2012. At the same time, a "greening" of the teaching force has occurred since the 1980s, as a quarter of all teachers now have five years of experience or less. Within the last 20 years, attrition among first-year teachers has increased by one-third, and 40 percent to 50 percent of all teachers leave within the first five years of entering the teaching profession.

Educator Accountability Reporting System (EARS) Since the 2006 LFC teacher preparation evaluation, institutions and the Legislative Education Study Committee created EARS to measure progress toward higher professional standards and financial support as required by Section 22-10A-19.2 NMSA 1978. While colleges of education continue to generate more revenue than is budgeted, this trend has lessened since the 2006 evaluation. New Mexico State University (NMSU), the University of New Mexico (UNM), and Western New Mexico University (Western) have increased the proportion of generated revenue that is allocated to colleges of education, while Central New Mexico Community College (CNM), Eastern New Mexico University (Eastern), and New Mexico Highlands University (Highlands) continue to allocate less than 50 percent of the revenue generated by education courses to their colleges of education. Among the state's institutions, colleges of education are large producers of student credit hours.

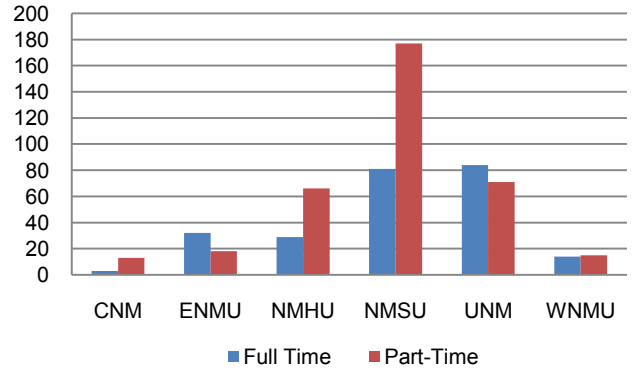
Table 2. College of Education Revenue and Expenditures, FY11

University	College Total Student Credit Hours (SCH)	College SCH as Percent of Institution Total	Adjusted Formula Revenue Generated by SCH (in thousands)	College Instructional Support Budget (with fringe benefits) (in thousands)	Expenditures per SCH (with fringe benefits)	Budget +/- Formula (in thousands)	% of Budget to Adjusted Formula Revenue
CNM	14,178	2%	$1,696	$715	$50	- $ 982	42%
Eastern	27,072	23%	$8,219	$3,805	$141	- $4,414	46%
NNMC	1,510	4%	$378	$576	$382	$199	153%
Highlands	20,652	25%	$7,161	$2,594	$126	-$4,567	36%
NMSU	48,373	11%	$15,847	$12,689	$262	-$3,158	80%
SFCC	4,035	4%	$449	$365	$90	-$84	81%
UNM	74,485	12%	$21,605	$16,068	$216	-$5,537	74%
Western	8,997	13%	$2,623	$1,826	$203	-$797	70%

Formula revenue generated is adjusted to exclude the 16.6 percent earmarked for the institution Source: 2011 EARS

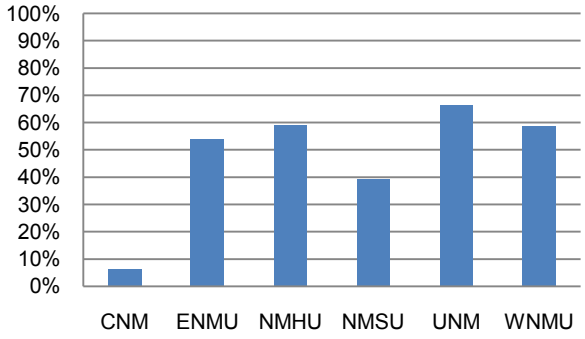
Since the 2006 LFC evaluation, the proportion of college of education faculty with doctorates has increased, though schools continue to rely on part-time faculty. Institutions tend to employ part-time faculty without doctorates to supervise clinical courses, and adjunct faculty are often current K-12 teachers.

Chart 1. College of Education Full-Time and Part-Time Faculty



Source: Colleges of Education

Chart 2. College of Education Faculty with Doctorates



Source: Colleges of Education

Faculty salaries have generally increased since the 2011 evaluation. However, in 2011, several of the colleges reported full-time entry-level salaries below statutory minimum salaries for level III teachers within New Mexico's three-tiered system.

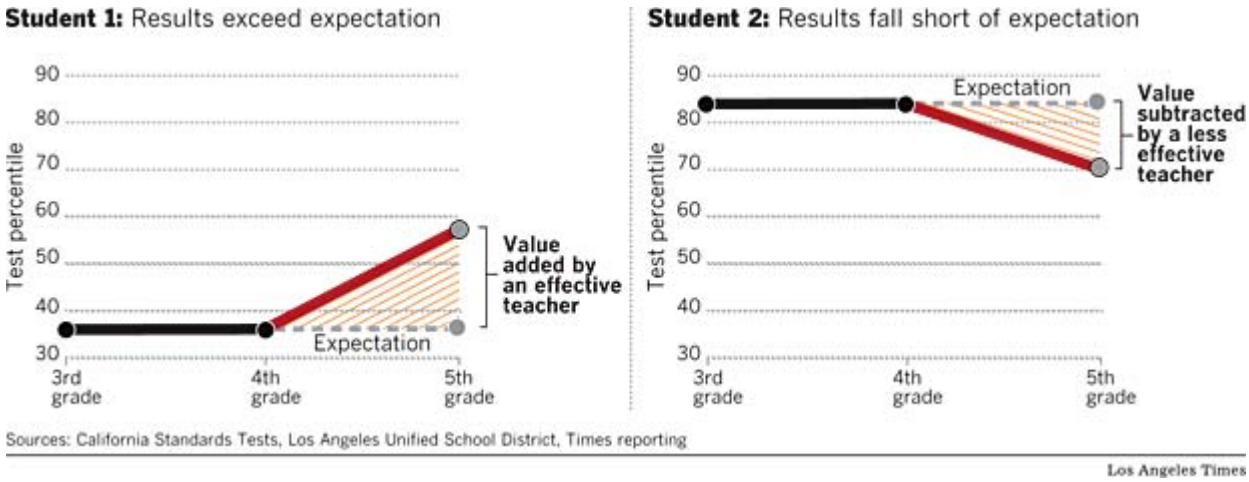
Table 3. Faculty Compensation SY10 - SY11

	Full-Time Faculty		Part-Time Faculty
University	Salary Range	Compensation per Course Range	Compensation per Course Range
CNM	$57,273 - $60,433	$ 5,727 - $6,043	$2,563 - $3,472
Eastern	$42,848 - $76,303	Salary only	$1,341 - $2,000
Highlands	$59,400 - $85,825	$7,425 - $10,659	$2,926 - $5,851
NMSU	$53,000 - $83,907	$6,625 - $10,375	$3,510 - $6,783
UNM	$54,825 - $130,549	$6,853 - $16,319	$2,714 - $6,000
Western	$44,159 - $63,367	Salary only	Salary only

Source: 2011 EARS

Value-Added Models. As has been done in many states and districts, New Mexico's standards-based assessments (SBA) can be used to calculate how much a teacher adds to student performance. While numerous approaches exist, in this evaluation, two years of prior SBA scaled scores as well as free- or reduced-price lunch (FRL) status were used to predict each students' reading and math scores for 2012 (**Appendix C**). The difference between that predicted value and the actual score, also known as a residual value, can be attributed to the influence of that student's teacher for SY12.

Figure 1. Calculating Residual Values



By averaging residual values for each student in a teacher's class for three years, the teacher receives a value-added score for a given school year. Some states and districts calculate these scores internally, while others, such as Tennessee, contract out the process.

LOW TEACHER ADMISSION REQUIREMENTS AND LICENSURE STANDARDS PERPETUATE LOW STUDENT PERFORMANCE

Despite investments in the state's three-tiered licensure system, colleges of education continue to attract and admit academically average candidates. In 2012, only 51 percent of New Mexico's students performed on grade-level in reading and only 43 percent performed on grade-level in math, as measured by the state's standards-based assessment (SBA). The three-tiered licensure system was a strategy to recruit and retain high-quality teachers, which in turn would help improve student achievement. However, admissions standards at New Mexico's colleges of education and the Public Education Department's licensure requirements have remained low.

Among New Mexico's five traditional licensing programs, schools maintain similar grade-point averages (GPA), applications, and coursework requirements for admission, though state law does not require minimum admission standards. All of the state's traditional preparation programs require a GPA between 2.5 and 3.0 for admission. While several universities maintain minimum ACT requirements for admission, none of New Mexico's colleges of education require minimum ACT scores. Additionally, most programs require applicants to successfully complete introductory coursework and general education courses within various content areas, such as math and English, prior to admission.

Table 4. Traditional Licensure Program Admission Requirements

University	Min. GPA	Min. Basic Skills Score	NMTA Content Test Completion	Min. ACT	Education Coursework	Content Coursework
Eastern undergraduate	2.8	240		17*	√	√
Eastern graduate	3.0	240		No	√	√
Highlands undergraduate	2.5	240		No	√	√
Highlands graduate	3	240		No	√	√
NMSU undergraduate	2.5	240	√	20*	√	√
NMSU graduate	3	240	√	No	√	√
UNM undergraduate	2.5	240		No	√	√
UNM graduate	3	240		No	√	√
Western undergraduate	2.5	240		21*	√	√
Western graduate	3	240		No		

*University admission requirement

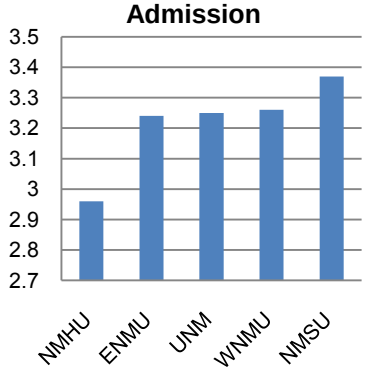
Source: 2011 EARS and 2011 Title II Reports

While alternative licensure programs generally maintain fewer specific admission requirements than traditional licensure programs, all alternative licensure programs in New Mexico require applicants to hold bachelor's degrees and pass the basic skills assessment. Five of the eight state-approved alternative licensure programs also require a minimum GPA for admission and most of the state's alternative programs require prior completion of university coursework within the licensure area.

ACT scores of candidates admitted to the state's colleges of education have not increased since the 2006 LFC evaluation. None of the state's colleges of education require minimum ACT scores for admission, unlike other schools within institutions, such as UNM's school of engineering, which requires a minimum math ACT subtest score of 25 and English ACT subtest score of 19. Statewide, admitted undergraduates tend to report slightly lower ACT scores, 20.1, than the average scores of graduate students, 21, and alternative licensure candidates, 20.2. At NMSU, ACT scores among undergraduates, 19.4, graduates, 19, and alternative licensure candidates, 17.8, all fall below the minimum ACT score of 20 required for undergraduate admission. While national research consistently

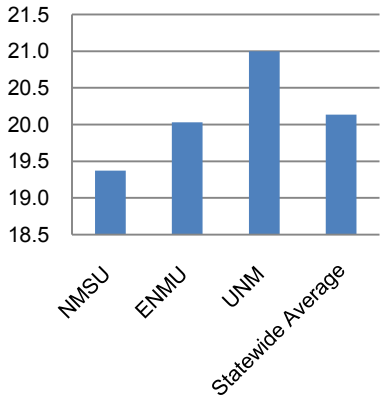
suggests colleges of education applicants tend to fall below the national average ACT score of 20, the average score of New Mexico teacher candidates, 20.1, is slightly above the state's overall average of 19.8. Trends in the GPAs of education students across the state are similar to ACT scores and align with K-12 student performance.

Chart 3. Average College GPA of Education Undergraduates at Admission



Source: 2011 EARS

Chart 4. Average Education Undergraduate ACT Score at Admission



Source: 2011 EARS
Data not available for Highlands or Western

Table 5. Mid-Range ACT Scores of All Admitted Undergraduates

University	ACT Score Mid-Range
Eastern	17- 23
NMSU	18 - 24
UNM	19 - 25

Source: The College Board

Establishing more stringent entrance requirements could improve prospective teacher effectiveness. Research demonstrates a correlation between teacher ACT scores and student reading achievement, though no significant impact on math was noted. A teacher with a record for high academic success adds about 4 percent to students' average academic achievement, an amount roughly equal to the impact of a single course on how to teach reading (Kennedy, Ahn, and Choi, 2008). In response, several states, including Colorado and North Carolina, have raised admission standards, including establishing minimum GPA requirements, requiring applicants to pass a pre-professional skills test in the top 75 percent, and requiring alternative licensure programs to adhere to the minimum admission requirements of traditional programs.

In addition, New Mexico programs do not meet standards developed by the National Council for Teacher Quality (NCTQ). The NCTQ recommends requiring teacher candidates to score in the top half of all college-going students on a test such as the ACT. The NCTQ also recommends a 3.0 GPA across a minimum of four college semesters and a minimum of a 3.0 GPA in the subject area to be taught.

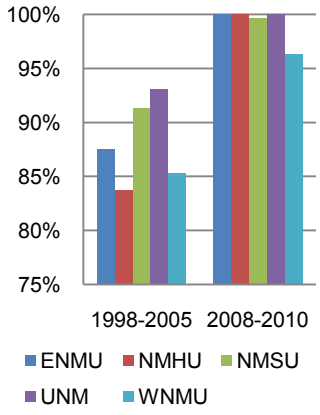
New Mexico's teacher competency exams provide little information about program quality as virtually all teachers pass. Similar to most states, New Mexico's licensure system requires the completion of a minimum of three competency examinations prior to level I licensure, including an assessment of basic skills, teacher competency, and content knowledge. Pearson Education, Inc. developed these tests, known as the New Mexico Teacher Assessments (NMTA).

All exams are scored on a scale of 300 points and administered six times each year. Teacher candidates first complete the basic skills assessment, designed to assess fundamental reading, writing, and mathematics skills generally acquired during middle school. With one exception, NMSU's alternative licensure program, all of the state's teacher preparation programs require teacher candidates to pass the basic skills assessment prior to admission.

To apply for a level I license, teachers must then pass the teacher competency assessment by licensure grade level, elementary or secondary, and pass a content area assessment, such as math, reading, or social studies. Beginning January 2013, Section 22-10A-7-(C) NMSA 1978 requires aspiring elementary teachers to pass an assessment of the science of teaching reading.

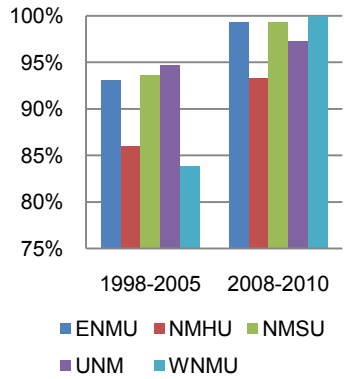
Since 2008, every institution's passage rates exceed 90 percent on all three tests. The 2006 LFC evaluation noted secondary competency pass rates lower than elementary pass rates, but secondary pass rates have since risen to within 3 percentage points of elementary rates. Also, low pass rates at Eastern and Highlands have increased since the 2006 LFC evaluation.

Chart 5. NMTA Basic Skills Pass Rates



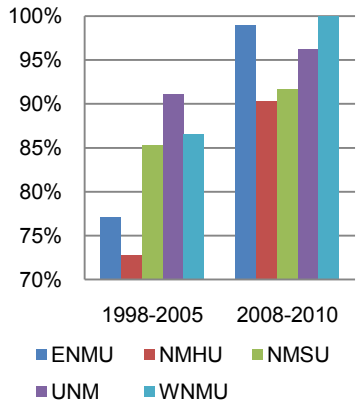
Source: LFC Analysis

Chart 6. NMTA Elementary Competency Pass Rates



Source: LFC Analysis

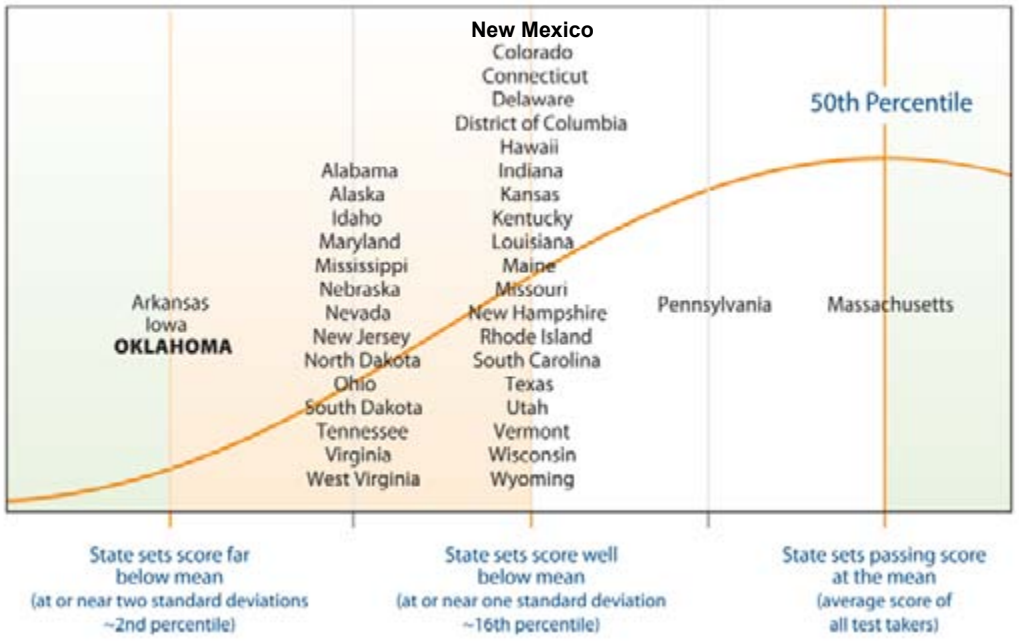
Chart 7. NMTA Secondary Competency Pass Rates



Source: LFC Analysis

NMTA cut scores do not effectively measure teacher quality. Since being set by the State Board of Education in 2000, the passing score for all NMTA assessments remains at 240, one standard deviation below the average score of 260. While research finds a teacher's content knowledge consistently predicts student performance, New Mexico's high passage rates mask these differences. While New Mexico's passage rates are similar to the 96 percent national passage rate in 2006, several states, including Massachusetts, Pennsylvania, and Tennessee, recently raised cut scores for their teacher competency exams.

Figure 2. State Teacher Exam Cut Scores



Source: National Council on Teacher Quality

Currently, teachers may retake the NMTA's an unlimited number of times. Of the 8,058 licensed teachers who passed the elementary content knowledge assessment between 2002 and 2012, 4 percent, or 348 failed at least one time, with 33 failing five or more times.

Table 6. Failed Elementary Assessment Attempts Before Passing 2002 – 2012

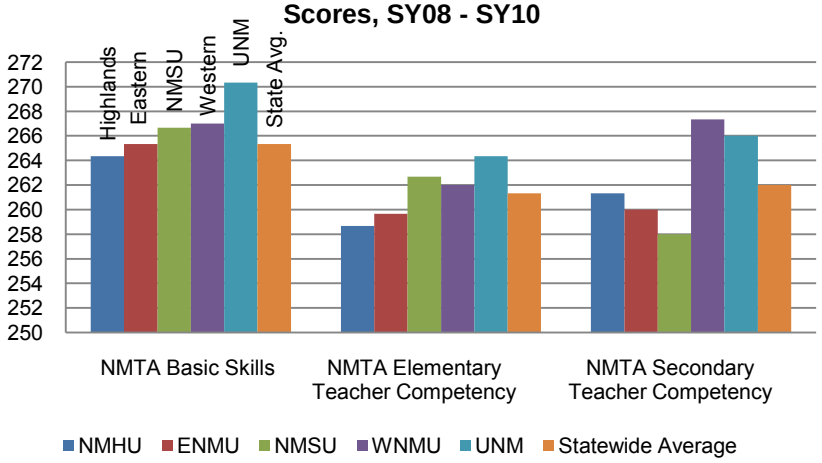
Number of Failures	Number of Teachers
1 - 5	326
6 - 10	19
11 - 17	3
Total	**348**

Source: LFC Analysis

Teachers who fail an NMTA at least one time perform lower than those who pass on their first attempt. For example, the average 2012 value-added score for teachers who failed the elementary content knowledge assessment at least one time, -0.2 points, is lower than the average for those who passed on their first attempt, 0.3 points.

NMTA score differences by institution follow the same trends as the differences in value-added scores. For admission, colleges of education require a passing score of 240 on the basic skills assessment, but higher scores indicate candidates more likely to be successful with K-12 students. Completers of UNM's traditional licensure program report the highest scores on the basic skills assessment, 270, and elementary competency assessment, 264, while Western completers report the highest score on the secondary competency assessment, 267. Highlands completers report the lowest basic skills and elementary competency scores, 264 and 259, while NMSU reports the lowest secondary competency score, 258.

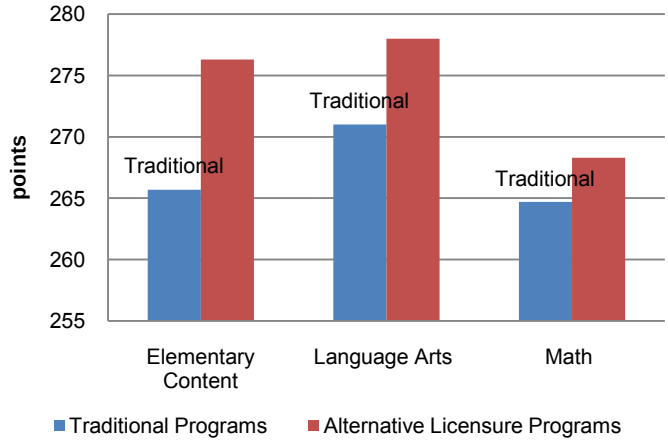
Chart 8. Traditional Preparation Program NMTA Scores, SY08 - SY10



Source: 2011 Title II Report

Pass rates and scaled scores are slightly higher among the state's alternative licensure completers. Overall, alternative licensure completers averaged a 278 on the basic skills assessment, compared with traditional completers earning 266; similarly, alternative licensure completers average 11 points higher on the elementary content exams, seven points higher on the language arts exams, and four points higher on the math content exams.

Chart 9. NMTA Content Knowledge Assessment Scores, SY08 - SY10



Source: 2011 Title II Report

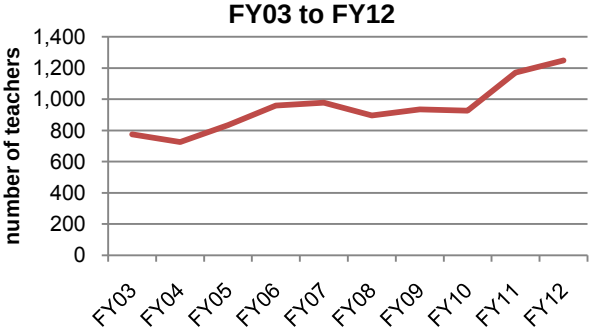
Teachers who score higher on the basic skills assessment, the elementary content knowledge assessment, and the mathematics content knowledge assessment tend to have higher value-added scores. Among the 1,365 teachers with 2012 value-added scores, scores on the basic skills, elementary content knowledge, and math content knowledge correlate to value-added scores. Raising cut scores for these assessments will likely correspond with increases in value-added scores, as teachers who earned a score of 260 on the math content assessment are predicted to add an average of 1.4 points to their students' SBA scaled scores compared with teachers who earned a minimum passing score of 240. Similar relationships exist between teachers' basic skills assessments and elementary content knowledge assessments.

The correlations between teacher assessment scores and value-added scores in New Mexico are consistent with national findings. According to education researcher Dan Goldhaber (2007), a standard deviation increase in teacher test performance corresponds to a 1 percent to 4 percent increase in student achievement. Similarly, the National Council on Teacher Quality recommends testing to confirm a teacher's content knowledge and pedagogical skills with the adoption of multiple rigorous content and pedagogical skills tests.

Raising cut scores would require higher performance from prospective teachers, although New Mexico's teaching supply can withstand increases to licensure standards. Since 2002, the average basic skills score for is 266, one standard deviation above the passing score of 240. Of the 19 thousand teachers with passing basic skills scores above 240 points, 4,349, or 23 percent, scored between 240 and 259. Similar trends exist for other elementary and secondary content assessments. Colleges of education will need to respond to higher NMTA standards by raising performance standards to ensure an adequate high-quality teacher pool.

New Mexico's teacher preparation programs currently supply an adequate number of completers to replace educators leaving the profession. In New Mexico, as is true nationally, teacher retirement rates appear to have peaked between 2011 and 2012. Based on Education Retirement Board data, 1,248 licensed New Mexico teachers, or 3 percent, retired in 2012, while LFC analysis predicts approximately 790 teachers will retire each of the next five years.

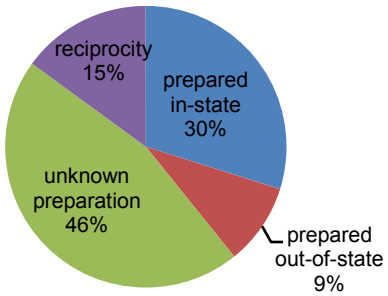
Chart 10. Teacher Retirement FY03 to FY12



Source: LFC Analysis of ERB Data

Overall, schools districts report 1,810 teachers left the workforce between SY11 and SY12, while New Mexico's colleges of education prepared 1,277 teacher candidates during SY10. Given that half of the state's teaching force is prepared in-state, this rate of preparation currently exceeds the need.

Chart 11. Preparation Institution of New Mexico's Licensed Teachers



Source: LFC Analysis of PED Data

Additionally, 26 thousand out of 47 thousand licensed teachers, or 56 percent, were not actively teaching during SY12, providing a sizeable eligible reserve of teachers.

Although New Mexico's overall supply of teachers is sufficient, particular content areas and geographic regions experience shortages. Specifically, the state and districts identify special education, math, science, and pre-K teachers as well as positions within certain rural communities as difficult to fill. Targeted incentives could be directed to these areas of need, while overall increases to minimum starting salaries could improve the state's ability to raise standards while attracting high-quality teaching candidates.

Recommendations

The Public Education Department should phase-in increases to the NMTA licensing cut scores, beginning in FY16.

The Legislature should couple increases in beginning teacher licensure standards with level I starting teacher salaries. To allow students and institutions to adjust for higher standards, the Legislature and PED should target implementation for FY16.

Colleges of education should raise admissions requirements, including the minimum NMTA basic skills assessment scores.

THE PUBLIC EDUCATION DEPARTMENT COULD BETTER OVERSEE PREPARATION PROGRAMS TO IMPROVE TEACHER QUALITY

The PED does not use student and teacher outcome data to approve and renew educator preparation programs. Current requirements for teacher preparation programs include 30 to 36 credit hours of professional education coursework, 24 to 26 credit hours in a teaching content area, and 14 weeks of field experience. Regulation limits alternative licensure coursework to no less than 12 credit hours and no more than 21 credit hours.

PED's approach to teacher preparation program approval and renewal relies heavily upon evaluations from the National Council for the Accreditation of Teacher Evaluation (NCATE). An advisory council of PED, the Professional Practices and Standards Council (PPSC), recommends renewal of preparation programs after reviewing NCATE reports. The educator preparation committee has met twice in the last year to approve several new programs, although the licensure committee has not met since 2007.

Currently, the NCATE accreditation standards PED relies upon focus on programmatic input measures, such as licensure exam pass rates and faculty qualifications (see **Appendix G**). Losing NCATE accreditation, however, does not correspond with loss of PED program approval, as Highlands remained on PED's list of approved teacher preparation programs in spite of losing NCATE accreditation between 2007 and 2012. Additionally, PED has not identified any institution as "at-risk" or "low-performing" for federal Title II reporting. Other states, including the 13 awarded Race to the Top funds, are linking student achievement to teachers and aggregating teacher effectiveness data to the preparation level.

The PED has the capacity to link student performance to teachers and colleges of education. Other states, including Tennessee, North Carolina, Texas, and Louisiana, use value-added outcome data to evaluate the effectiveness of their colleges of education, and federal reporting will soon likely require the same approach. Ohio uses measures of teacher effectiveness within their higher education performance-based funding formula.

Of the 21 thousand teachers with active classroom assignments in New Mexico, the LFC used five years of student data to determine value-added scores for 1,365 teachers in SY12 (**Appendix C)**. For the 1,365 teachers with student data from SY10, SY11, and SY12, the average value-added score is 0.3 points, meaning these teachers helped their students score 0.3 scaled score points above the students' predicted scores.

Interpreting Value-Added Scores

Given the low proficiency rates across the state, moving large numbers of students to grade-level performance requires significant gains. For example, students scoring at beginning steps, the lowest level, need to increase scaled scores by at least 10 points to be considered at grade-level. Even making two points of progress per year, it will take such a student five years to become proficient.

Table 7. Statewide Value-Added Scores, 2012

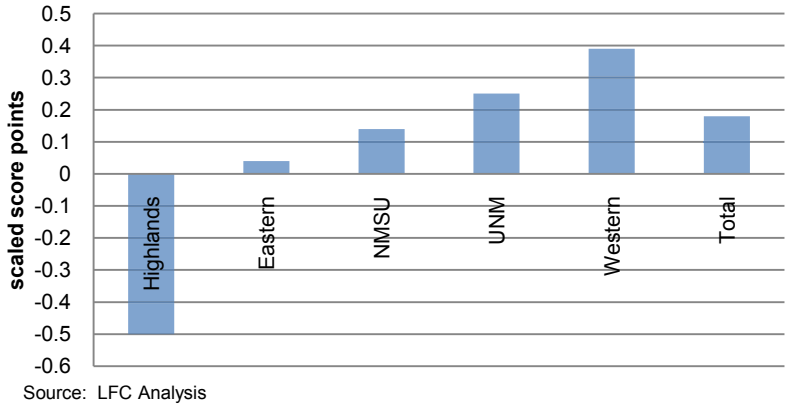
	Number of Teachers	Minimum	Maximum	Average	Std. Deviation
Mean Residual 2010	2,556	-8.6	8.2	0.1	2.2
Mean Residual 2011	2,484	-8.5	7.9	0.1	2.1
Mean Residual 2012	3,459	-9.4	10	0.2	2.0
Value-Added Score, 2010 - 2012	1,365	-5.7	7.4	**0.3**	1.7

Source: LFC Analysis

Average value-added scores by college range from -0.5 points to 0.4 points, indicating need for overall improvement to increase student achievement. Of teachers with less than eight years of experience, those from Eastern, NMSU, UNM, and Western add value to their students' performance, while those from Highlands average a negative value-added score. The average value-added score for these teachers prepared in-state, 0.18 points, is

nearly identical to the average of 0.16 points for teachers prepared out-of-state. Given the state's current proficiency rates, however, making "catch-up growth" will require higher value-added scores across New Mexico's colleges of education.

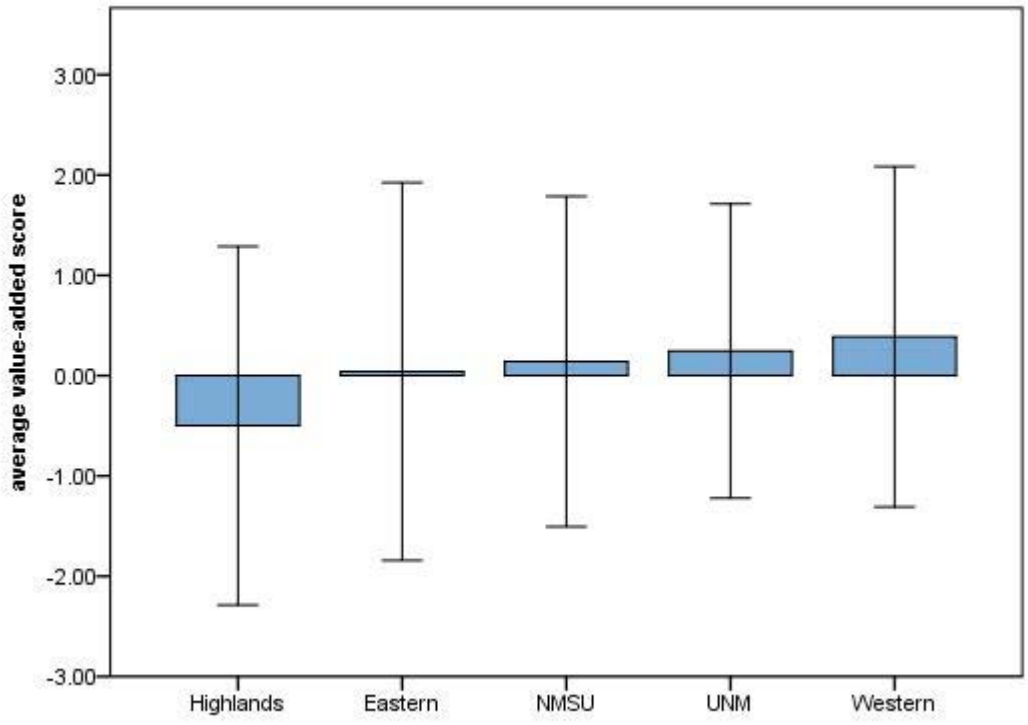
Chart 12. Average Value-Added Score, 2012



Source: LFC Analysis

Within each college, however, performance varies widely, resulting in significant overlap between schools. For example, while the average difference between Highlands and Western is 0.9 points, the range at Highlands is from -2.3 to 1.39 compared with Western's range of -1.3 to 2.1.

Chart 13. Range in Value-Added Scores between College, 2012

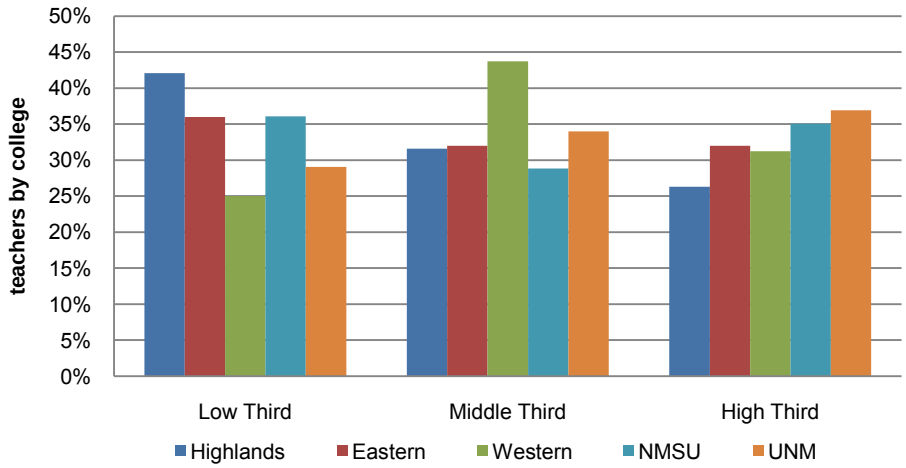


Source: LFC Analysis

Error Bars: +/- 1 SD

To interpret these differences, some states, such as Tennessee, compare colleges by ranking teachers into performance groups. When the 548 teachers in this analysis are similarly sorted into thirds, the distribution is unequal. At UNM, for example, 37 percent of teachers perform in the highest third, compared with 26 percent of Highlands's teachers; also, Western has a higher percentage of teachers performing in the middle third, 44 percent, than at either the low end, 25 percent, or the high end, 31 percent.

Chart 14. Value-Added Distribution by College, 2012



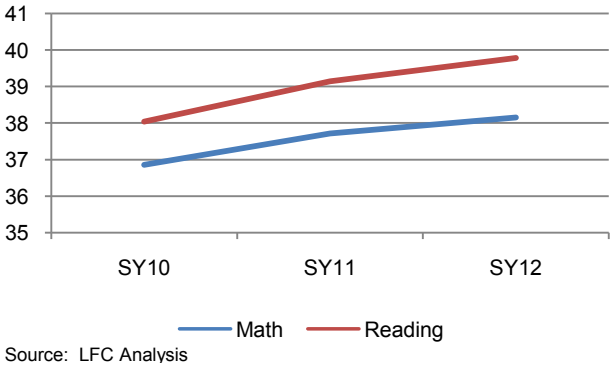
Source: LFC Analysis

These differences between colleges highlight the importance of carefully selecting candidates, raising licensure standards, improving program quality, and creating incentives within the higher education funding formula.

On average, alternatively licensed teachers' value-added scores are higher than traditionally licensed teachers. The average value-added score for an alternatively licensed teacher in New Mexico is 0.4, compared with an average value-added score for traditionally licensed teachers of 0.3. In 2012, 11 percent, or 3,173 of the teachers licensed in New Mexico, completed alternative programs, which allow candidates who have already earned a bachelor's degree to earn a teaching certificate by completing coursework in how to teach.

Student gains in scaled scores also highlight differences between programs. In SY12, the average SBA scaled reading score for all students was 39.8, with 40 considered proficient, while the average scaled math score for all students was 38.2.

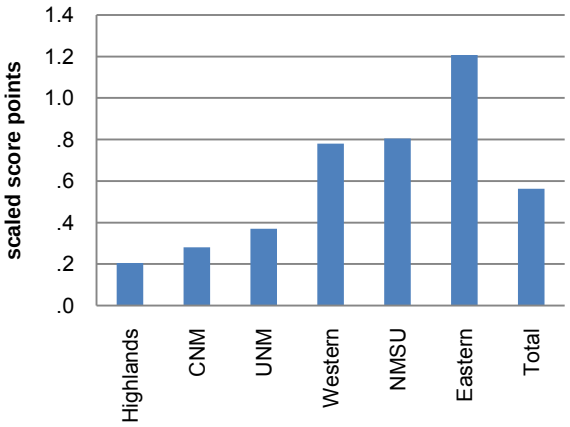
Chart 15. Average SBA Scaled Scores, Teacher Prep Cohort



Source: LFC Analysis

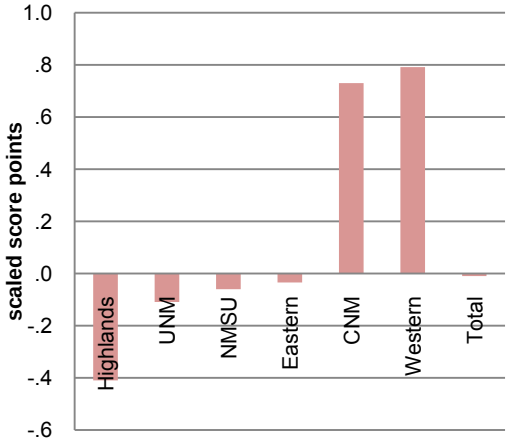
Year-to-year, changes in scaled scores indicate relative growth, with the same score from one year to the next representing one expected year of growth. From SY11 to SY12, Eastern prepared teachers whose students made the greatest average scaled score gains in reading, 1.2 points, while Western prepared teachers whose students made the greatest average scaled score gains in math, 0.8 points. Highlands had the lowest average gains in reading, 0.2 points, as well as math, -0.4 points.

Chart 16. Average Reading Scaled Score Gains, SY11 to SY12



Source: LFC analysis

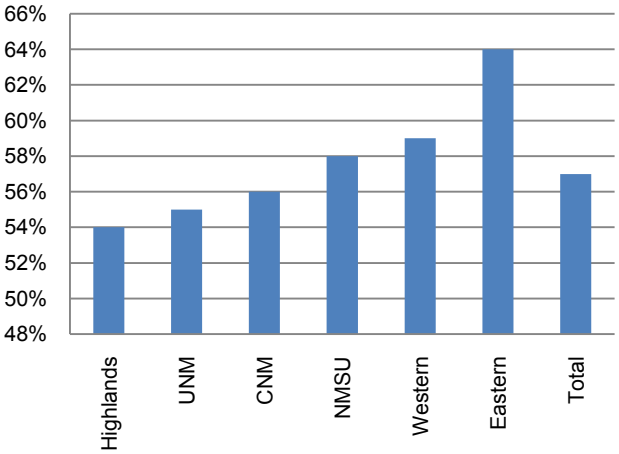
Chart 17. Average Math Scaled Score Gains SY11 to SY12



Source: LFC Analysis

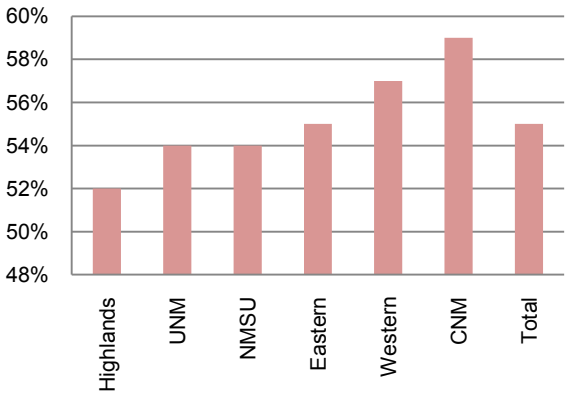
The percentage of students who made one year's worth of growth by school shows similar trends: Eastern has the highest percentage in reading, 64 percent; CNM has the highest percentage in math, 59 percent; and Highlands has the lowest percentages in reading, 54 percent, and math, 52 percent. Statewide, 57 percent of students grew by at least one year in reading and 55 percent grew by at least one year in math.

Chart 18. Students with at Least One Year of Reading Growth SY11 to SY12



Source: LFC Analysis

Chart 19. Students with at Least One Year of Math Growth SY11 to SY12



Source: LFC Analysis

Practitioners and employers agree about recent program completers' areas of weakness, many of which could be better-addressed through coursework. According to LFC survey data, teachers report feeling least prepared to meet the needs of students with disabilities, teach English language learners (ELL), and effectively use student data (**Appendix E**). These reflections are important because self-perceptions of effectiveness often drive decisions to stay in the profession (Kee, 2012).

Additionally, principals agree preparation is weakest in the same three areas that teachers identified, though principals reported traditional program completers are more adequately prepared than completers of alternative licensure programs. Principal agreement tended to be highest for NMSU completers and lowest for alternative licensure completers.

Table 8. Principals Who Agree Teachers are Well or Sufficiently Prepared to Meet Teacher Expectations

University	Manage the Classroom	Teach Reading	Teach Math	Support Students with Disabilities	Teach English Language Learners	Use Student Data
Eastern	77%	88%	86%	76%	72%	73%
Highlands	74%	78%	77%	64%	74%	67%
NMSU	81%	91%	90%	79%	79%	75%
UNM	77%	86%	89%	73%	73%	74%
Western	75%	87%	88%	67%	73%	71%
Alternative Licensure	43%	57%	59%	38%	38%	53%

Source: LFC Survey

Educator preparation programs generally fail to meet standards of high quality regarding data and assessment preparation, but a few programs demonstrate rigorous and authentic preparation. Research-based best practices call for teachers to frequently assess students, analyze data, and adjust instructional strategies to drive student achievement. While colleges of education should integrate data analysis into coursework, an LFC review of course syllabi suggests New Mexico's teacher preparation programs do not fully meet the best practices outlined in the National Council of Teaching Quality's *Linking Assessment and Instruction Innovative Configuration*. Often, teacher-candidates completing special education licensure programs receive more extensive preparation to use student data than teachers preparing for elementary or secondary licensure.

However, several colleges better prepare students to use data. Western, for example, requires all traditional teacher candidates to complete an assessment course, while Eastern's blended elementary and special education program serves as a model of rigorous preparation in data-driven instructional practices because teacher candidates perform several diagnostic assessments, analyze results, and develop intervention strategies accordingly.

Table 9. College of Education Data and Assessment Coursework

Criteria	CNM	Eastern	Highlands	NMSU	UNM	Western
Course Devoted to Data/ Assessment	√	√*	√*			√
Technical Topics Related to Data and Assessment		√*	√*	√*		√
Types of Assessments		√ *	√*	√ *		√
Issues Related to Assessing Diverse Populations		√ *	√*	√		√*
Teacher Candidates Design Assessment	√	√	√	√	√	√
Candidates Conduct a Diagnostic Assessment	√*	√	√*		√*	√*
Candidates Analyze Student Work	√	√	√	√	√	√
Candidates Analyze Student Data Over Time	√	√	√*	√	√*	√
Candidates Analyze Student SBA Data			√*			√

* included in courses not required for all programs Source: LFC Analysis of Syllabi Provided by Colleges of Education

Special education teachers are most likely to receive extensive preparation in classroom management, and several programs devote more time to developing teachers' classroom management skills. Classroom management plays a crucial role in student achievement and can significantly influence the persistence of novice teachers in the profession (Ingersoll and Smith, 2003). Based on an LFC review of course syllabi for traditional licensure programs, coursework falls short of the practices outlined by the National Council of Teacher Quality's *Classroom Organization and Behavior Management Innovation Configuration*. Only Western and CNM require all teacher candidates to complete a classroom management course. Other programs primarily address classroom management through reflection during field experiences, a potentially research-based practice.

Table 10. Classroom and Behavior Management Coursework

Criteria	CNM	Eastern	Highlands	NMSU	UNM	Western
Classroom/ Behavior Management Course	√	√	√ *	√*		√
Curriculum Addresses Classroom Environment	√	√	√ *	√		√
Curriculum Addresses Conveying Expectations	√	√	√ *	√*	√	√
Curriculum Addresses Behavior Reduction Strategies	√	√	√ *	√	√	√
Teacher Candidates Develop a Classroom Management Plan		√ *	√ *	√ *	√	√

*Coursework not required for all programs Source: LFC Analysis of Syllabi Provided by Institutions

Programs generally prepare teacher candidates to serve the needs of English language learners (ELL) and other exceptional populations, but special education candidates have more opportunities to apply these skills. Previous LFC evaluations highlighted the achievement gaps observed among New Mexico's ELL and special education students, reflecting the challenges teachers face improving educational outcomes for these populations. All of New Mexico's traditional licensure programs require general education teacher candidates to complete an introductory special education course, but few purposefully integrate special education coursework with fieldwork practices. Eastern, however, has blended its elementary and special education programs so candidates complete fieldwork to practice teaching in multiple settings, and Western's special education course includes a fieldwork component.

Table 11. ELL and Special Education Courses and Activities

Criteria	CNM	Eastern	Highlands	NMSU	UNM	Western
ELL Course Required		√	√*		√*	√
Curriculum Includes Characteristics and Research Related to ELL Students	√	√	√*	√	√*	√
Candidates Learn and Practice ELL Strategies	√	√*	√*	√	√*	√
Fieldwork Ensures Work with ELL Students		√	√*	√*	√*	
SPED Class Required		√	√	√	√	√
Curriculum Includes Characteristics and Research Related to Students with Disabilities		√	√	√	√	√
Candidates Learn SPED Strategies and Accommodations	√	√*	√	√	√	√
Fieldwork Ensures Work with Students with Disabilities		√*	√	√*		√

*Coursework not required for all programs (elementary and secondary). Courses required only for Teaching English as a Second Language or special education licensing programs not counted in this matrix.

Source: LFC Analysis of Syllabi Provided by Institutions

Few New Mexico teacher preparation programs require candidates to take a class in how to teach English language learners, and elementary teachers are more likely than secondary teachers to complete such a course. Several universities require teacher candidates to complete multicultural education coursework, but these courses focus upon issues of diversity and social justice rather than the characteristics of language acquisition or strategies that

support ELL students. Colleges of education often integrate strategies for serving ELL students by requiring candidates to detail modifications in lesson plans. Western provides a model for promising ELL preparation, as all teaching licensure candidates complete a multicultural education course and an ELL methods course.

A newly implemented reading exam is intended to measure teachers' readiness in the science of reading instruction. New Mexico's School Personnel Act requires teachers seeking an elementary or special education license to complete six credit hours of reading methods coursework and teachers seeking a secondary license to complete three hours. According to the state's 2010 *Study Reading Curricula in Teacher Education, HJM16*:

- Despite wide variance in program quality, every program showed room for improvement in one or more areas;
- Many New Mexico teacher education programs "missed the target in addressing the science of reading instruction to a disappointing degree"; and
- New Mexico should rigorously assess teacher candidate knowledge of how to teach reading through an examination.

New Mexico's colleges of education have since changed reading methods curricula and beginning in January 2013, elementary teacher licensure candidates must pass a rigorous reading assessment. Results of this assessment will provide additional evidence about the quality of reading methods courses.

High quality fieldwork produces positive student outcomes. Student teaching is funded between $133 and $635 per credit hour, depending on the course level, with student teaching coursework generating $1.7 million in funding formula revenue in SY11. Research shows first-year teachers who graduate from programs actively involved in selecting field placements, with minimum experience levels for cooperating teachers, and requiring supervisors to observe student teachers at least five times have higher student achievement than those whose field experiences do not meet these criteria. Other research-based field experiences practices include the following:

- Require teacher candidates to demonstrate beginning teacher competence prior to student-teaching placement;
- Integrate fieldwork throughout the preparation curriculum;
- Place field experience students in high-poverty, high-performing school placements;
- Provide field experience students with written and oral feedback opportunities after frequent observations by clinical faculty;
- Provide year-long student-teaching experiences; and
- Evaluate teacher candidates based on student learning data (Boyd et al, 2009).

These practices require greater oversight and rigor than the standards detailed by the National Council for Accreditation for Teacher Education (NCATE), which all New Mexico colleges of education currently hold. Student achievement data as well as feedback from practicing educators suggests existing fieldwork experiences are insufficient.

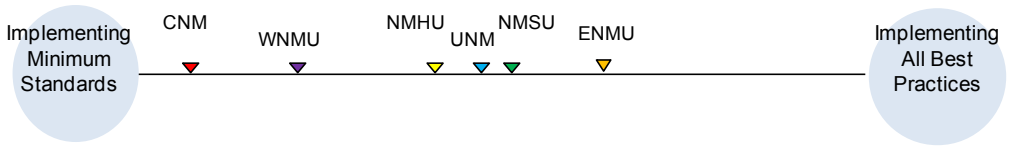
Teachers and practitioners consistently rank field experiences as crucial in the development of novice educators. According to an LFC survey of over 200 principals, 80 percent strongly agreed that student teaching is a critical element of teacher preparation, and 86 percent strongly agreed that strategies for effective classroom management, often practiced through student-teaching, are critical. Principals tended to rate student teaching as more critical than content knowledge (**Appendix E**).

Similarly, the 4,000 teachers surveyed referenced student-teaching and hands-on fieldwork as the courses that most prepared them for success.

> *"Actually teaching in the classroom is what provided me with the best preparation – courses helped and provided some theoretical background, but it was the practice of teaching that did it."*
> Teacher response from LFC survey

Though New Mexico's teacher preparation programs generally exceed minimum field experience standards articulated in administrative rule and NCATE accreditation, schools fall short of fully implementing research-based best practices. At Eastern, Highlands, and UNM, cooperating teachers must meet minimum experience requirements prior to serving as supervisors, and Eastern and UNM student-teachers appear to receive more frequent, structured observations and debriefing sessions with faculty supervisors and cooperating teachers than candidates in other programs. However, candidates are not always placed in professional-development school settings, and placement within clinical school sites often do not persist throughout fieldwork courses (see **Appendix F** for the scoring rubric and supporting research). Additionally, student-teaching structure varies among alternative licensure programs because teacher candidates often teach full-time while completing coursework.

Chart 20. Progress Toward Implementing Student Teaching Research-Based Best-Practices



Source: LFC Analysis

While several colleges of education have adopted site-based models, research suggests some models are more effective than others. Eastern, NMSU, and UNM, for example, have moved all or parts of fieldwork courses to public school sites, providing clinical settings for practicum coursework. This involves closer collaboration with districts and schools, but these models generally do not persist throughout fieldwork or are not available to all teacher candidates. One example of a promising site-based model is UNM's partnership with Bandelier Elementary. UNM integrates fieldwork at Bandelier to provide rigorous and meaningful experiences for teacher candidates. This model is unique because of the extensive collaboration between Bandelier Elementary and UNM, the number of student-teachers at the site, continuous teacher-candidate placement within a single school site, and selective practicum placement.

In SY12, the SBA math and reading gains of Bandelier's fifth grade students, all co-taught by UNM student-teachers, were significantly higher than other fifth graders in Albuquerque Public Schools (APS) and the state.

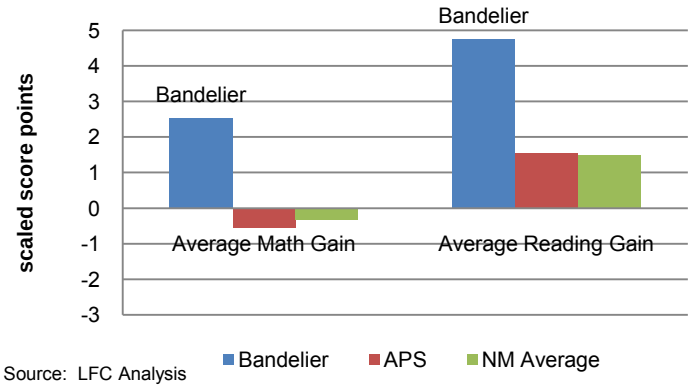
Table 12. Bandelier Fifth Grade SBA Gains, SY12

5th Grade Cohort	Students Who Grew in Math	% Who Grew in Math	Students Who Grew in Reading	% Who Grew in Reading
Bandelier	37	64%	46	78%
APS	3,072	57%	2,919	57%

Source: LFC Analysis

While students in APS grew 1.6 scaled score points in reading and lost 0.6 scaled score points in math between their fourth- and fifth-grade years, fifth-grade students at Bandelier Elementary grew an average of 4.8 scaled score points in reading and 2.5 scaled score points in math.

Chart 21. Bandelier Fifth-Grade Scaled Score Increases, SY12



Source: LFC Analysis

UNM also reports Bandelier student-teachers experience less "praxis shock," or the feeling of being underprepared that many new first-year teachers report. Finally, placement rates of Bandelier teacher-candidates suggest program completers possess the skills principals seek in new teachers. Ten of 12, or 83 percent, of UNM students in the first Bandelier cohort were immediately hired, compared with the first-year placement rate of 44 percent among all newly licensed teachers in 2011.

Recommendations

The Public Education Department, with the colleges of education, the LFC, and the Legislative Education Study Committee, should develop a methodology for calculating average value-added scores by institution, calculate this value-added score annually, and identify performance benchmarks for each college of education.

The Public Education Department's Professional Practices and Standards Council should review student outcome data and educator retention data to supplement NCATE institutional reports in the program approval and renewal process.

The Higher Education Department should discontinue funding programs that lose state approval.

The Higher Education Department should identify options for incorporating teacher preparation program outcome data and employment retention rates in the higher education performance-based funding formula through the funding formula task force.

Colleges of education should improve and expand research-based teacher clinical experiences for traditional licensure programs, including:

- cluster student teachers at high-poverty, high-performing sites;
- require student-teacher candidates to complete a selective placement process demonstrating basic teacher competencies prior to student-teaching approval;
- select mentor teachers with demonstrated records of student achievement;
- offer on-site instruction and professional development for all staff at student-teaching sites;
- require a minimum of five formal student-teaching observations coupled with opportunities for feedback from supervising faculty; and
- adopt co-teaching strategies.

INCREASING ENTRANCE STANDARDS, EXIT STANDARDS, AND PROGRAMMATIC QUALITY WILL RAISE ADMINISTRATOR QUALITY

Admission standards and licensure requirements are not preparing school leaders with the greatest potential. State law does not establish admission requirements for administrative licensure programs, though research suggests that recruitment and selection are central components in the program design of highly effective school leadership programs (Darling-Hammond et al., 2007). Principal preparation also matters, leading New Mexico's colleges of education to adopt and streamline coursework to align with the *Interstate School Leaders Licensure Consortium* (ISLLC) leadership standards in 2009. As a result, the core courses completed by principal candidates are similar across programs, though considerable qualitative differences in administrative internships exist. While regulation requires only that administrators complete 180-hour internship over the course of a year, research suggests internship quality, particularly a residency model, plays a key role in the development of school leaders.

In New Mexico, administrator programs generally maintain low admission requirements. Admission practices could better identify candidates by relying on recommendations that strategically identify candidates with leadership potential. Currently, only UNM and Western require recommendations from a supervisor or individual who can discuss the candidate's leadership potential. Also, selection focuses on years of teaching experience, rather than measures of instructional effectiveness described in previous LFC evaluations.

Eastern and UNM require a level II license, while Highlands does not specify years of teaching experience or licensure requirements for admission. All of the state's administrative licensure programs require a 3.0 GPA for admission.

Table 13. Administrator Preparation Program Admission Requirements

University	GPA	Minimum Years Teaching Experience	Licensure Level	Other Requirements (recommendations, essays, resume)
Eastern	3.0	6	II	√
Highlands	3.0			√
NMSU	3.0	3		√
UNM	3.0	4	II or III	√
Western	3.0			√

Source: 2011 EARS

Administrator licensure requirements limit the supply of highly qualified school leaders. Obtaining an administrative license in New Mexico requires a minimum of six years teaching experience or seven years for out-of-state applicants. In contrast, Texas and Oklahoma require only two years and Colorado and Arizona each require three years.

By the time candidates are eligible for administrative licensure, they earn more per day as level III teachers than as an entry-level principal. Based on typical contract lengths for each position and the statutory minimum annual salaries of $50 thousand for level III teachers and $60 thousand for elementary principals, level III teachers earn a minimum of $278 per day compared with $273 per day for elementary principals. Opportunities for administrative licensure earlier in an educator's career would lessen these pay differentials.

Between 2008 and 2010, 100 percent of all administrator program completers passed the administrator assessment. In addition to level III licensure, administrator candidates must pass the educational administrator assessment, which is also developed by Pearson Education, Inc. and has a cut score of 240 out of 300. Pass rates and scaled scores for Highlands and NMSU, the two largest producers the state's administrators, were missing from the 2011 Title II report.

Table 14. New Mexico Educational Administrator Assessment Pass Rates, 2008 - 2010

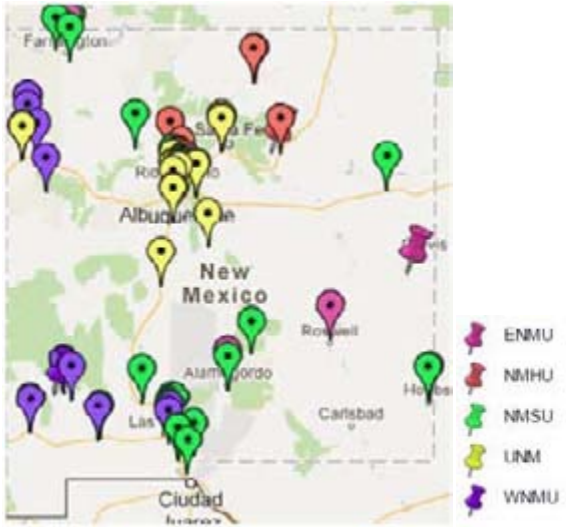
University	First-time Pass Rate	Average Scaled Score
Eastern	100%	NR
Highlands	NR	NR
NMSU	NR	NR
UNM	100%	271
Western	100%	264.5
Statewide	**100%**	**262.3**

Source: 2011 Title II Report

As measured by school grades, differences in the quality of principal preparation are minimized when student poverty is taken into account. While New Mexico's school grading system allows principal effectiveness comparisons, after controlling for student poverty, most of the differences in preparation programs even out (see **Appendix D** for a description of the principal population and methods for this analysis).

Principals tend to serve in communities surrounding the college that prepared them for school leadership. The geographic nature of principal placement leads certain administrator programs to produce candidates who tend to serve in areas with higher levels of poverty than others.

Figure 3. Placement of Principals Prepared by New Mexico Institutions



Source: LFC Analysis

Principals prepared by Western, in particular, tend to serve in schools with higher levels of poverty than principals prepared by other administrator preparation programs in the state.

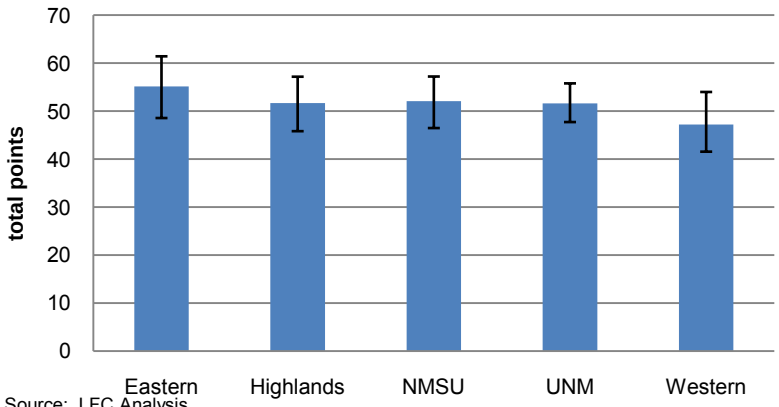
Table 15. FRL Levels by Principal Preparation Program, 2011

	Principals in Schools with <50% FRL	Principals in Schools with 50-75% FRL	Principals in Schools with 75-100% FRL	Average FRL
Eastern	2	2	4	66%
Highlands	4	3	9	68%
NMSU	4	10	8	68%
UNM	11	11	11	63%
Western	0	2	17	82%
Statewide Total				66%

Source: LFC Analysis

Differences in SY12 school grade totals attributed to administrator preparation programs exist but are less meaningful when poverty is taken into account. When comparing schools' total grade values and student growth values, statistically significant differences appear between programs. For example, Western's principals have lower total school grade scores, 47.2 points, than principals prepared by other in-state programs, 53.9 points.

Chart 22. Average School Grade Total by Administrator Preparation Institution Before Adjusting for Poverty



Source: LFC Analysis
Error bars represent 95% confidence intervals

Source: PED

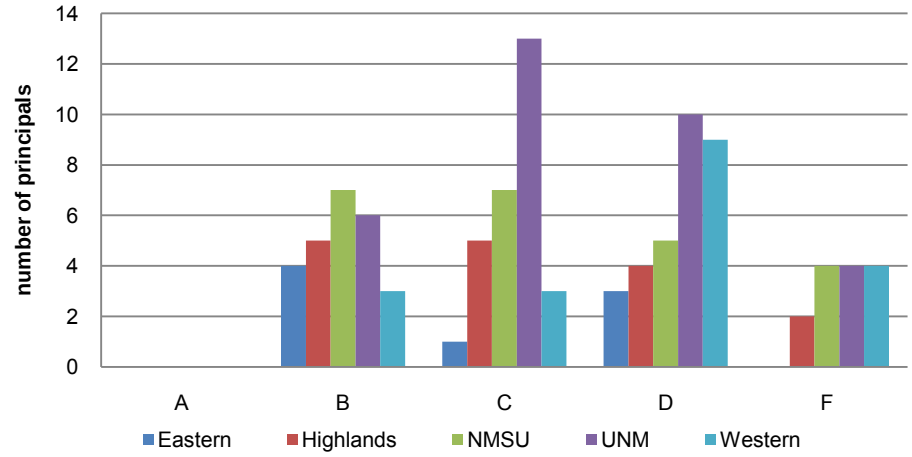
Table 15. School Grades

Total School Grade Points	Letter Grade
75.0 - 100.0	A
60.0 - 74.9	B
50.0 - 59.9	C
37.5 - 49.9	D
0.0 - 37.4	F

Source: PED

Of the sampled principals, administrators from Highlands and Eastern tend to serve schools with higher school grade totals; 63 percent of the principals associated with each school earned B's or C's, whereas 68 percent of the principals prepared by Western serve at schools earning D's or F's. However, principals from administrator preparation programs with lower school grade totals also serve in schools with higher poverty levels.

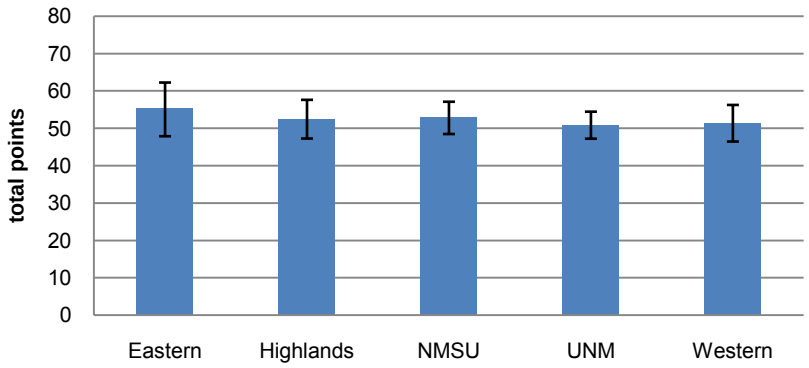
Chart 23. Distribution of School Grade Totals Among Principal Sample



Source: LFC Analysis

After controlling for school poverty levels, school grade differences attributed to administrator programs shrink. The adjusted school grades reported below estimate a college's average total school grade at the state's average poverty level of 66 percent. Even after controlling the effect of poverty level on school grades, school grade-point values sorted by preparation program differ, but these estimates overlap among colleges and are quite small.

Chart 24. School Grade Totals Among Administrator Preparation Institutions After Adjusting for Poverty

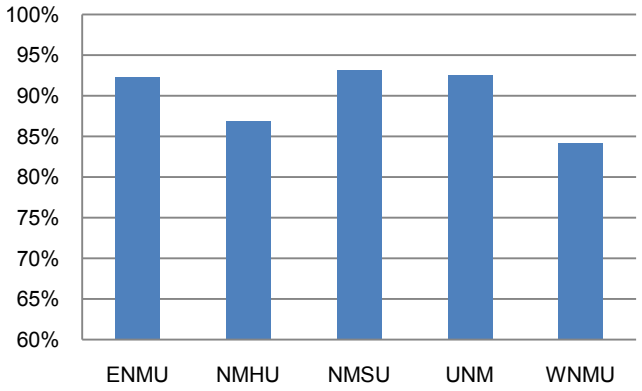


Source: LFC Analysis

This trend, the reduction of significant differences after controlling for the effects of poverty, is also true for sub-categories within school grades, including current status and growth of both high-performing and low-performing students. However, the relationship between administrator preparation colleges and the growth of a school's top three student quartiles is statistically significant for elementary principals.

Despite the overlap in school performance, practitioners and district administrators perceive school leader preparation programs differently. Based on an LFC survey of New Mexico's administrators, principals from NMSU, UNM, and Eastern report the highest levels of preparation (**Appendix E)**.

Chart 25. Principals Who Agree They Were Well or Sufficiently Prepared Overall

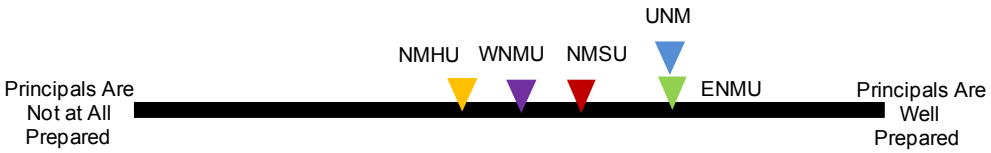


Source: LFC Survey

District administrators rate UNM and Eastern graduates as best prepared overall among principals prepared in-state, but the survey did not show significant differences within specific competencies.

Figure 4. Average Level of District Administrator Response to the Following Question:

"What is your overall evaluation of how well the institution prepares principals to effectively lead schools?"



Source: LFC Survey

Survey data tended to mirror trends observed in the analysis of student outcome data. Schools with higher ratings tended to produce greater student growth, and schools with lower ratings tended to produce less student growth.

UNM's principal preparation partnership with APS is a promising clinical practice worth replicating across the state. The Alliance of Leading and Learning (ALL) is a new principal preparation partnership between UNM, APS, and the New Mexico School Leadership Institute (NMSLI). Among the 13 members of the first cohort, 12 gained administrative licensure and ten are now employed as assistant principals or deans in high-need schools. This placement rate is seven times higher than the statewide rate of 12 percent in 2012. Recent research funded by the Wallace Foundation supports aspects of ALL, including a careful selection process; full-time, semester-long residencies; and follow-up mentoring.

The Alliance of Leading and Learning can guide improvements among other programs. UNM, APS, and the NMSLI developed this federal grant-funded partnership to improve student success by carefully selecting principal candidates, identifying administrative mentors with records of student success, and matching these mentors with principal candidates. APS administrators co-teach all coursework with university faculty. Co-teachers receive grant-funded stipends, and their instruction enables future principals to connect theory with practice. After coursework, principal candidates complete a semester-long, full-time internship alongside mentor principals. APS provides long-term substitutes to fill the classroom positions of these principal interns at a cost of $9,700 per

candidate. This approach starkly contrasts other schools of education that have moved toward entirely online internships in which interns complete logs documenting activities while maintaining full-time positions.

The program's most significant costs are operational and mentorship support, including the salaries of a program manager, district mentor principal, administrative assistant, and NMSLI staff. Staff plan to track the program's success by measuring completers' administrator retention rates and could also track school performance for each completer.

Recommendations

The Public Education Department should raise licensure cut scores for administrators.

The Public Education Department should link public school grades to administrator preparation institutions and consider this data during administrator program approval and renewal.

Colleges of education should improve and expand research-based administrator clinical experiences, including:

- strategically recruit and select principal candidates with the greatest leadership potential;
- require full-time, semester-long residency for principals; and
- partner with districts to develop and support principal residency and mentoring programs.

The Legislature should reduce minimum teaching requirements to obtain an administrative license.

NEW MEXICO'S EDUCATOR REPORTING SYSTEM CAN BE SIMPLIFIED AND IMPROVED BY INCLUDING OUTCOMES DATA

The educator accountability reporting system (EARS), designed to provide the state with information about program performance, primarily includes inputs that overlap with federal reports. In response to the 2006 LFC evaluation of teacher preparation programs, the state initiated the educator accountability reporting system to provide an annual update of how well colleges are preparing educators from pre-entry to post-graduation.

Expanded with data on administrator preparation, the EARS report is to include demographic and performance characteristics of students and program completers, hiring and retention data, and financial measures. While statute requires PED and colleges of education to collaborate to develop the EARS report, only colleges of education have undertaken this task.

EARS data replicates information included in federal Title II reports. Although much of the EARS report is similar to information annually submitted to the U.S. Department of Education, differing data definitions require institutions to recalculate the same measures. Colleges of education consider the process redundant and burdensome, and the PED does not appear to rely upon EARS to assess how well the state is preparing educators.

EARS repeatedly generates the same findings, but no progress has been made to address concerns or collect teacher persistence and student outcome data. Though statute requires inclusion of educator retention rates and student outcome indicators, EARS does not because colleges of education lack access to this data. Other recurrent EARS findings include:

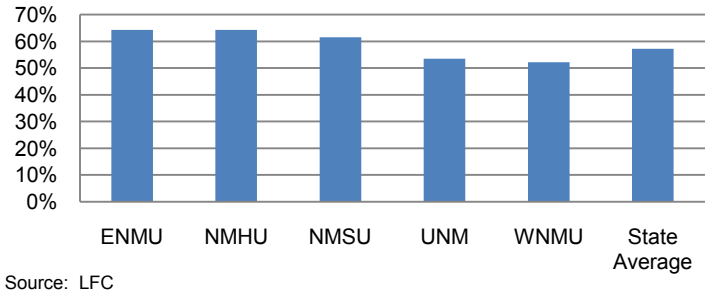
- Information that should be reported into the student teacher accountability reporting system (STARS), such as teacher and administrator preparation institute, either cannot be reported into STARS because the appropriate data fields are missing, or information is inaccurately reported and never verified;
- Teacher and administrators self-report preparation institutions when completing New Mexico Teacher Assessments, producing errors in scores and pass rates attributed to colleges of education; and
- Financial data does not accurately capture the contribution of colleges of arts and sciences, which provide much of the general education content instruction to teacher candidates.

Title II reports will likely soon require colleges of education to report student outcome data as well as information about teacher retention, the same data EARS does not include.

PED reporting on employment retention will encourage the colleges of education to increase the percentage of teachers who stay in the profession for at least three years. Based on LFC analysis of PED data, among teachers prepared in-state and licensed in 2008 and 2009, an average of 57 percent still taught three years later, indicating turnover in the first three years among New Mexico teachers is higher than the national average of 25 percent.

Teacher persistence rates at Eastern, Highlands, and NMSU exceed the state averages. At 64 percent, Highlands and Eastern have the highest three-year persistence rates among newly licensed teachers.

Chart 26. Average Three-Year Persistence Rate of Teachers Licensed in 2008 and 2009



Source: LFC

On average, teachers who left the classroom between 2011 and 2012 had a value-added score of -0.01 points, while teachers who remained in the classroom had value-added scores of 0.3 points. While currently New Mexico teachers who leave the workforce are slightly less effective than those who remain, aggregating teacher retention data and student achievement data to the institution level could lead to improvements in both outcomes.

Recommendations

The Legislature should revise statute to substitute the federal Title II report for the educator accountability reporting system, including student outcome and teacher retention data by college.

The Public Education Department should annually calculate a three-year employment retention rate for each college of education.

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

HANNA SKANDERA
SECRETARY OF EDUCATION

SUSANA MARTINEZ
GOVERNOR

November 30, 2012

Mr. David Abbey, Director
Legislative Finance Committee
325 Don Gaspar, Suite 101
Santa Fe, NM 87501

RE: Teacher Preparation Programs

Dear Director Abbey:

Thank you for the opportunity to respond to the draft evaluation on Teacher and Administrator Preparation in New Mexico. Please accept my compliments to your staff for their professionalism and collaborative approach throughout the evaluation process. The Public Education Department (PED) is committed to providing a rigorous and effective framework for the improvement of Teacher Preparation programs.

The current evaluation of teacher preparation programs appears to be thorough and objective and points to a number of issues that will help us establish a more effective teacher workforce that is capable of ensuring preK-12 students receive the education they need to excel in the 21st century. We are pleased that the evaluation has accounted for multiple sources of information from within the state of New Mexico, and has also used ongoing, well-researched, areas and best practices- regarding teacher preparation- that are occurring on a national level.

A key component of the LFC evaluation is the use of student achievement data to evaluate "early career" teachers and their impact on student outcomes. The present study supports the better understanding of the programs that are providing the quality rigor and relevance to pre-service teachers.

The exit conference between LFC and PED was held Tuesday, November 20, 2012 and the draft report was discussed. The department does not have any recommended changes at this time. We look forward to working together as we move toward establishing pre-service criteria that are robust, fair and truly focus on improving the teaching skills of all candidates.

Thank you again for the opportunity to comment on the evaluation.

Warm regards,

Hanna Skandera
Secretary-Designate
Public Education Department

HS/mm

LFC CLARIFICATION OF ISSUES IDENTIFIED BY THE NEW MEXICO DEANS AND DIRECTORS OF EDUCATION

The New Mexico Deans and Directors response identifies concerns regarding the methods used to determine value-added scores. Other states conduct similar analysis to evaluate the performance of colleges of education and this evaluation drew from those methods. As described in the LFC's recommendations, calculating these scores raised methodological questions that will need to be addressed by the Public Education Department, but the results presented in this evaluation provide reliable and valid insight into overall performance.

Teacher Population Selection. The LFC identified teacher and administrator preparation institutions based on data files provided by the Public Education Department. In an attempt to improve the quality of this data, the LFC sent these preliminary lists to each of the six institutions for verification but did not receive responses from all institutions. Based on feedback from Eastern, Western, Highlands, UNM, and CNM, discrepancies in completer status were sent to the Higher Education Department for verification and final lists were compiled. Following the analysis, the colleges of education did not notify the LFC of which students were misidentified. As the state calculates value-added scores by institution, the PED will need to work with the HED and the institutions to accurately connect program completers to the appropriate college of education.

Teacher Population Selection: Time Since Degree. The LFC used years of experience, not graduation data, as the selection criteria because variables beyond preparation institution are likely to increasingly influence a teacher's effectiveness over time. The specific number of years, eight, was chosen to maximize the number of teachers included in the analysis. Additionally, the LFC received completion dates only from Highlands, and based on that data, all but two of the teachers in the sample completed degrees by 2004, eight years prior to SY12. Of those two, one completed in 2003 and the other in 1997 and both have less than eight years of experience; removing these two cases lowers the average value-added score for Highlands. When calculating value-added scores by institutions, the PED will need to work with the HED and the institutions to define an acceptable range for years of experience and date of completion.

Accuracy of VAM for Estimating Teacher Effects. As noted in this report, previous LFC evaluations have highlighted limitations and cautions regarding the use of student data to measure teacher effectiveness. Most critiques emphasize student data should be one amongst multiple measures used to assess teacher effectiveness. Similarly, this evaluation recommends considering student data along with other outcomes indicators to more accurately and completely measure the performance of New Mexico's colleges of education.

Failure to Clearly Describe Statistical Analyses. Technical details are provided in Appendix C and Appendix D. Additionally, the LFC responded to individual methodological inquiries. Regarding differences in average value-added scores between institutions, Chart 13 illustrates the wide range in scores between institutions and resulting overlap.

Response to: Teacher and Administrator Preparation in New Mexico.
Report #: 12-13. Date: 12/5/2012

Prepared by: New Mexico Deans and Directors of Education

Forward

This document is in response to the findings and recommendations of the LFC report on teacher and administrative preparation in New Mexico. While we believe that the process used was done with the best of intentions, the conclusions and recommendations go far past what the findings would indicate. There appear to be a number of critical problems with the value added modeling (VAM) methodology, the use of the National Council on Teacher Quality (NCTQ) as an indicator of best practices, and the findings regarding field experiences. In addition, there are other issues that illustrate the contention that the conclusions are not supported by the available data.

Methodological Issues

1. Teacher Population Selection Inaccuracies

The selection of teachers from each institution for the Value-Added Model (VAM) analysis part of the study is seriously flawed. The LFC analysts' data was not verified independently with the actual degree award records of each institution with varying degrees of inaccurate attributions of teachers-to-institution resulting. For instance, in the following institutions, the reported number of teachers holding degrees from an institution are contrasted with an exhaustive internal records search from that same institution:

NMHU: Of the 19 students selected for the NMHU analysis: (a) three received graduate degrees but not undergraduate degrees; (b) two students graduated between 2000 and 2005; (c) three were not found or unverified in the NMHU system. Overall, only 84% of the NMHU teachers were accurately identified by the LFC audit.

NMSU: NMSU received a teaching list with 247 names and social security numbers. However, of the 247, 67 came with SSNs and there were no student records found based on the SSN. Further, 9 on the list had majors outside of the COE for a total of 76 unverified students. Overall, only 69% of the NMSU teachers were accurately identified by the LFC audit.

UNM:

- 203 Student SSNs were given to UNM by the LFC, with 201 of these having a UNM record (defined as minimally taking one course)
- 180 received a degree from the College of Education (89%), some of them more than one degree.
- Overall there was a 59% match rate for administrators and an 89% match rate for teachers.

2. Teacher Population Selection: Time Since Degree

The LFC audit report contends that their VAM analysis was performed using teachers with 8, or fewer, years of experience. However, the institutional records indicate that this is not the case. There has been no discussion of how long a teacher preparation program is responsible for their graduated teachers. Even the selection of 8 years or less is completely arbitrary and without foundation. The question that must be answered is, "At what point does a teacher's life experiences and subsequent learning disqualify them from being attributed and thus evaluated by any given institution of higher education?"

NMHU: Of the 19 students used in the audit, only 8 received an undergraduate degree or completed their licensure requirements since the fall of 2005 (8 years). Three of the students received their degrees prior to 2000. One graduated in 1985 and two in the 1990s. Overall, only 42% of the sample completed their preparation programs at 8 or fewer years.

NMSU: Of those students who graduated from NMSU, 15% of the list had graduated between 1988-2006. Overall, only 85% of the sample completed their programs at 8 or fewer years.

UNM:

- Correctly identified degrees that go back to 1983.
- 11 COE degrees were awarded to these students in the 1980's.
- 34 COE degrees were awarded to these students in the 1990's.
- The remainder in the 2000's.
- Overall, only 78% of the sample completed their programs at 8 or fewer years.

3. Best Practices Reference (NCTQ)

The use of the NCTQ as a indicator of "best practices" in teacher preparation (pages 10, 14 and 22) is ill-considered and without basis in fact. The National Council of Teacher Quality (NCTQ) is not a government agency, is not sanctioned by federal or state government or by higher education accreditation associations. NCTQ is a privately funded advocacy group that conducts superficial studies of colleges of education that do not meet the most minimal standards of good research. The studies consist of NCTQ requesting documents (e.g., course syllabi, resumes of full and part time faculty, program handbooks, rubrics for culminating projects, etc.) from colleges of education. The contents of the documents are evaluated against NCTQ standards. To date, NCTQ's standards have not been independently vetted by experts in educational research. There is no verification of any data, nor is there an opportunity for the colleges to respond or correct misinterpretations.

4. Controversy Regarding the Accuracy of VAM for Estimating Teacher Effects

The Value Added Model (VAM) has many advocates, however numerous researchers have criticized the use of VAM for rendering inaccurate results. The issues raised in the sampling methods (above) exacerbate the final values which are used extensively by the LFC to draw conclusions and make recommendations about teacher and administrator preparation programs.

The use of VAM to estimate teacher effects on student achievement is controversial and this should be acknowledged in the LFC report. One of the primary problems with VAM is that teacher influence on student achievement cannot be easily distinguished from other student variables. A quote from a recent review of VAM makes this point:

"The default assumption in the value-added literature is that teacher effects are a fixed construct that is independent of the context of teaching (e.g., types of courses, student demographic compositions in a class, and so on) and stable across time. Our empirical exploration of teacher effectiveness rankings across different courses and years suggested that this assumption is not consistent with reality. In particular, the fact that an individual student's learning gain is heavily dependent upon who else is in his or her class, apart from the teacher, raises questions about our ability to isolate a teacher's effect on an individual student's learning, no matter how sophisticated the statistical model might be." (p. 18)*

* Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010) Value-Added Modeling of Teacher Effectiveness: An exploration of stability across models and contexts.

Educational Policy Analysis Archives, 18 (23). Retrieved [date], from: http://epaa.asu.edu/ojs/article/view/810

5. Failure to Clearly Describe Statistical Analyses Used to Identify Institutional Differences

While the LFC report describes differences among New Mexico Schools and Colleges of Education on a number of dimensions, the report does not consistently describe the statistics used to determine if these are "real" differences or differences that might be occurring by chance due to things like small sample sizes, which can skew the results of a study like the LFV report. Consequently, it is possible that the LFC report is inaccurately describing differences in performance among the institutions that don't actually exist, which is misleading and should not serve as the basis for policy decisions.

Other Issues with LFC Recommendations

Increasing Field Experiences

One of the LFC report recommendations is that, "Colleges of education should improve and expand research-based teacher clinical experiences for traditional licensure programs…(p.26)." However, there is no relationship between student credit hour production (SCH) and funding at UNM, or most universities. Most systems use a historical budgeting model that does not fund by SCH - and so it does not matter how much, or how little, we produce in terms of our budget. We realize that this may not be how the LFC might look at budgeting, but it is the reality in the university.

The LFC's estimated revenue (SCH funds) associated with Field Experiences across all the state's institutions was $1.7 million dollars SY11 (page 27). This value does not reflect the scope or cost of actually placing, monitoring, and supervising student teachers in the field. UNM's cost alone for field services is approximately $1.2 million dollars per year, or 71% of the SCH funds generated by the Field Experiences courses across all teacher preparation programs in New Mexico. These costs include: honorarium for cooperating teachers in the schools, supervision budgets for college personnel, administrative costs, and travel. Student teaching and other field experiences are quite expensive - and they are unavoidable. It is simplistic to think that funds recovered from SCH even get close to the real cost of these events. The Deans and Directors are adamantly opposed to the imposition of any additional calls for more field experiences until the full cost of these activities are completely understood and become part of a realistic funding model.

Requiring more stringent admission requirements to college of education programs, e.g., higher ACT scores and higher minimum NMTA basic skills assessment scores.

According to the LFC report, "…colleges of education continue to attract and admit academically average candidates.…"(p.10) even though the average GPA for admission to teacher preparation programs of ten colleges of education is 2.78. The LFC recommendation is for colleges of education to "establish more stringent entrance requirements [that] **could** improve prospective teacher effectiveness" (p.14), because "…a teacher with a record for high academic success adds about 4 percent to students' average academic achievement" (p.14). The LFC cited other schools within institutions requiring minimum ACT math and English subtest scores that are higher than the minimum ACT score required for undergraduate admission to the university. The LFC report repeatedly focused on ACT scores as a possible admission requirement, even though the studies reviewed by the LFC found "…no significant impact on [student math] achievement …."(p.14) The Deans and Directors of Colleges of Education are not opposed to higher admission standards, but live with the reality that many incoming freshmen that are academically superior candidates gravitate to schools of engineering and other sciences. This is not unexpected and can easily

be attributed to starting salaries. Specifically, the median starting salary for engineering majors graduating in 2012 is $59,000 (www.forbes.com) whereas entry level teachers in New Mexico earn an annual salary of $30,000. "Academically average candidates" being drawn to education is not a college of education admission issue, it's a state and national issue that colleges of educations are burdened with and somehow still manage to overcome by providing the state with excellent teachers.

Raising admissions requirements by increasing the minimum NMTA basic skills assessment scores. The LFC report cites research that indicates an increase in teacher test performance "…corresponds to a 1 to 4 percent increase in student achievement…."(p. 18). Further, the LFC report indicates that NM's teacher preparation programs currently supplies an adequate number of completers with the "…rate of preparation currently exceeding the need (p.18)". This analysis is flawed as evidenced by the teacher deficit cited in the LFC's own analysis. The LFC states that, "school districts report that 1,810 teachers left the workforce between SY11 and SY12, while New Mexico's colleges of education report 1,277 teacher candidates completed licensure preparation programs…(p.15)" This leaves a deficit of 533 teachers not available to the state and the statement that the supply is sufficient to demand is not supported.

The LFC recommendation goes so far as to say "…New Mexico's teaching supply can withstand increases in licensure standards…(p.18)" yet notes that "…particular content areas and geographic regions experience shortages….special education, math, science, and pre-K teachers….(p.19)".

The Deans and Directors are not opposed to higher admission standards, and are anxious for the results of increasing the minimum NMTA basic skills assessment scores in Massachusetts, Pennsylvania, and Tennessee before implementing this projected solution to solving the academic achievement gap.

Colleges of education should improve and expand research-based teacher clinical experiences, specifically incorporating field experience in high-poverty, high-performing schools, place students in professional-development schools, select mentor teachers with demonstrated records of student achievement, and offering on-site instruction and professional development for all staff at student teaching sites.

A very small percentage of schools in New Mexico's 89 public school districts would qualify as "high-poverty-, high-performing schools" and those schools are not geographically accessible to every college of education. The LFC report does not define a "professional-development school", identify professional development schools or provide clarity what aspects of a professional development schools contribute to the success of entry level teachers or increased student achievement. For colleges of education to continue to produce a surplus of teachers, faculty from colleges of education must develop cooperative and respectful relationships with district and school-site leadership, and the expectation of only accepting mentor teachers with demonstrated record of student achievement is impossible and short-sighted, especially since strategies for managing the classroom are the most desired qualities of entry level teachers (p. 52). Finally, expecting colleges of education to offer professional development for all staff at student teaching sites is another unfunded fiscal and resource burden for the colleges of education. Several of the colleges of education place students at multiple sites across the state, and due to budget constraints rely heavily on part-time faculty to deliver the necessary courses, provide field supervision, and coordinate placement of teacher candidates.

In Summary:
The Deans and Directors of the College of Educations across the state of New Mexico continuously strive for program improvement. They do so by sharing information; mentoring new members; regularly meeting to confer, and remaining committed to developing passionate, competent and capable entry-level teachers and administrators. We look forward to the next challenge that has the potential to truly impact teacher and administrator preparation programs. And as always, we appreciate the opportunity to present our position which acknowledges the need for continual improvement to our programs while simultaneously asserting the quality of those same programs.

Response to LFC Audit Presented December 5, 2012
Prepared by
Michael A. Morehead
Dean College of Education
New Mexico State University

First I would like to thank Michael and Rachel for their openness and willingness to share their report with the deans and directors. This willingness to listen to our questions and concerns is greatly appreciated. I support their recommendation for the teacher education programs to work in concert with PED to develop a better system for determining the quality of the teacher education programs in New Mexico. Additionally, many of the assessment tools used by Rachel and Michael are being discussed at the national level. These strategies might be used by the Department of Education to determine teacher education quality in the states. Therefore this study may give us a snapshot of what the future could bring.

You will find additional information on the research we have conducted on Value Added Models and Teacher Retention Research. I believe this research strongly suggests that VAMs and retention data are not reliable and valid methods to assess the quality of teacher education programs. These studies have been provided to the LFC staff and are available to you.

My primary concern involves the extensive use of standardized tests to imply or make assumptions about student growth, quality teaching and then quality teacher education programs. Throughout the report, whether it be ACT scores as a basis for determining quality of candidates entering teacher education, or VAM scores to determine the ability of teachers to impact learning, all findings on student learning/achievement are based on some type of standardized test.

Determining quality teaching and improved student achievement using VAM scores, higher ACT scores, and increasing the pass rate on licensure tests makes for erroneous and misguided assumptions and findings related to quality teaching. **The circular illogical argument that higher ACT and NMTA scores of teachers leads to a higher standardized test score for students which then translates into better student achievement misrepresents the definition of student achievement and provides a very limiting definition of student achievement.**

It is my view that the country and the state of NM have been misled by the accountability movement, because of its overuse and misuse of standardized tests. Every state, national and international ranking has some linkage to standardized testing. We often are told that the United States is ranked 20th in math or 25th in science on international tests. However no one tells you that when the international comparison is with similar students who take the test that we are number one in the world or in the top five on most tests. **The major factor that impacts students' performance on standardized tests is poverty and the economic status of the family**.

Let me say with certainty that standardized tests for any group does not and cannot give educators a true picture of a student's learning, knowing or academic achievement. Standardized tests only give us a snapshot of where a child or an individual are on a continuum specifically designed by the testing company.

In the LFC report, it is suggested that by requiring higher scores on teacher licensure NMTA exams there will be a correlated positive impact on student learning and achievement in schools. **(Again an erroneous finding, because test scores do not and cannot give us a total picture of student learning and achievement)**. The chart on page 12 of the LFC report suggests that NM licensure cut scores are low and should be raised, and thus there would be higher test scores in NM for children. NM has a higher NMTA cut score requirement for teachers than North and South Dakota and Iowa. Using LFC logic NM students should have higher tests scores than student in those states. It is obvious that reasons other than NMTA teacher licensure scores must be impacting the test scores in the aforementioned states. Again trying to link NMTA scores to future teacher quality and program quality is misguided and imparts an inaccurate view of education graduates and programs.

The perspective presented in the LFC report suggests that standardized test scores demonstrate student achievement. I contend that this misrepresents what real learning and teaching is about. In addition, rating quality teacher education programs using VAM scores also misrepresents teacher education graduates' true impact on student learning. Tests produced by a national company that are standardized cannot and do not accurately inform us about student achievement and learning. A one day snapshot is an unfair way to judge something as complicated as student achievement.

It is illogical to base the success of a profession and students on standardized test scores. Our country has been misled by the accountability movement's attempt to take a simple example, such as a test score, as the primary rationale to judge a very complex and multifaceted profession.

Central New Mexico Community College

December 4, 2012

CNM Response to LFC Teacher Preparation in New Mexico

We would like to take this opportunity to thank the LFC program evaluators who lead us through this process. Although we believe that there are flaws in the data collection and analysis in this report, we appreciate that this program evaluation granted access to data that has been previously inaccessible to higher education. The lack of access and transparency highlights the need for a P-20 data system in the state that would allow access to data on program graduates for use in making programmatic decisions. As the principal alternative licensure program in the report, we would like to respond to this evaluation from an alternative licensure perspective.

Admissions & Exit from Program

In order to qualify for alternative licensure, a candidate must hold a Bachelor's degree or higher and have 24-30 credits in their core content area. This requirement ensures that a candidate enters the program with prerequisite content knowledge in their field and meets state licensure requirements. We have higher numbers of teachers graduating in high need areas such as secondary math, science and special education due to the recruitment of careered individual who come to the teaching profession with invaluable life and work experience. Students are held to high standards throughout the program and are only allowed to participate in their final supervised student teaching with a GPA of 3.5 of higher in their coursework. Students must demonstrate proficiency in all of the New Mexico teacher competencies in order to successfully complete the program.

Field Experience

The report recommends that all programs provide an intensive year-long student teaching assignment in a high-poverty high performing Highlandss exemplary practice, but the reality of alternative licensure is that many students are currently working in the field or hold full time jobs that do not allow for a one-year intensive field experience. In addition, alternative licensure students are not eligible for financial aid, unlike traditional education students, which makes taking a year off of work to complete student teaching financially unfeasible.

To compensate for the limited number of credit hours and time constraints faced by alternative licensure students, field-based assignments are required in all coursework in addition to a final semester of supervised student teaching. In order to participate in the final supervised field experience, all students must have a GPA of 3.5 in their coursework, which demonstrates proficiency in lesson plan development, assessment, classroom management, and reading instruction.

Part of our mission is to serve area schools by recruiting qualified teachers in high need areas and provide support for their staffing needs. To do this, we work with schools in many different areas of Albuquerque and are continuously building relationships with schools serving high poverty communities as placement opportunities for our students.

525 Buena Vista SE
Albuquerque, NM 87106-4096

Central New Mexico Community College

Assessment and Evaluation

CNM currently has one class in General Education (elementary/secondary), *Curriculum Development Assessment and Evaluation,* and one class in Special Education, *Methods and Materials for Special Education,* that address curriculum and classroom-based assessment. Based on our needs analysis and feedback from students and administrators currently in the field, as of Fall 2013 we will be requiring two courses in Curriculum Development Assessment and Evaluation for General Education and Special Education. We strongly believe that these changes will help CNM teachers address the needs of their students through data-driven decision making.

Students with Disabilities and ELL

Currently, state law limits the number of credit hours an alternative licensure program can offer to 12-21 credit hours. This limitation does not allow us to offer a separate course for general education students related to special education or working with English language learners (ELLs) as recommended in the report.
In order to address the needs of the students, every course in the general education program has content related to working with students with disabilities and ELLs.

In addition, CNM offers topics courses for working with students with disabilities and additional coursework that can lead to an endorsement in Teaching English to Speakers of Other Languages (TESOL). We encourage all of our students to take this coursework in addition to their program requirements.

Alternative licensure creates a pathway to teaching that brings diverse candidates and teachers qualified to teach in high need areas. The benefits these candidates bring to the profession of teaching outweigh the limitations imposed by restricted credit hours and as a program we are constantly striving to improve education for all students in New Mexico by providing the highest quality teacher preparation.

525 Buena Vista SE
Albuquerque, NM 87106-4096

APPENDIX A: Project Information

Evaluation Objectives

1. Follow-up on the 2006 LFC evaluation of teacher education programs.
2. Analyze the relationship between teacher and administrator education programs in New Mexico and student performance as measured by New Mexico's standards-based assessments.
3. Review the status of New Mexico's educator accountability reporting system (EARS).

Evaluation Procedures

- Reviewed best practices in teacher and administrator preparation, including the 2009 National Council on Teacher Quality evaluation of New Mexico's teacher education programs.
- Reviewed the relationship between performance data, including standards-based assessment scaled scores and employment retention rates, and teacher and administrator preparation programs.
- Interviewed and electronically surveyed faculty and staff from New Mexico's colleges of education and currently practicing teachers, principals, and district administrators.
- Reviewed applicable laws and regulations; LFC file documents, including the 2006 evaluation of teacher preparation programs; relevant performance reviews from other states; and performance measures.

Evaluation Team

Michael Weinberg, Lead Program Evaluator
Rachel Mercer-Smith, Program Evaluator

Authority for Evaluation

LFC is authorized under the provisions of 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. LFC is also authorized to make recommendations for change to the Legislature. In furtherance of its statutory responsibility, LFC may conduct inquiries into specific transactions affecting the operating policies and cost of governmental units and their compliance with state laws.

Exit Conferences. The contents of this report were discussed with the Public Education Department on November 20, 2012 and the Deans of the Colleges of Education on November 16, 2012.

Report Distribution. This report is intended for the information of the Office of the Governor; the Public Education Department; the Higher Education Department; New Mexico's Colleges of Education; 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.

Charles Sallee
Deputy Director for Program Evaluation

APPENDIX B: Public Education Department Report Card

Performance Overview: The strategic elements considered to evaluate the effectiveness of public schools are student achievement, teacher quality, and student persistence. Between FY06 and FY12, student performance as measured by the percent of students scoring proficient or above on the New Mexico Standards-Based Assessment (NMSBA) increased 10.4 percentage points in math but decreased 6 percentage points in reading. Statewide data from the FY12 assessment shows modest improvements of one percentage point in math and reading compared to FY11. Based on FY12 assessment data, 49.2 percent of students scored below proficient in reading and 57.1 percent students scored below proficient in math. While overall proficiency rates are showing incremental increases, proficiency rates for certain grades and subjects are below FY11 rates. For example, third graders reading at or above proficiency decreased 0.5 percentage points from FY11, and have decreased 5 percentage points since FY10.

The Public Education Department (PED) notes a decrease from 67.3 percent to 63 percent in FY11's four-year cohort graduation rate. Listed subgroups (students with disabilities, economically disadvantaged, Caucasian, American Indian, African American, etc.) did not improve over FY10. Part of the decrease is attributed to a new calculation that captures students not historically included in the calculation; however, it is unclear what portion of the decrease is a result of the new calculation.

For FY12, the department did not calculate adequate yearly progress (AYP); however, the department estimated that had it been calculated, approximately 98 percent, or 811 schools would have failed to make AYP. The state implemented a new accountability system that gives schools a letter grade between A and F based largely on student performance on the New Mexico standards-based assessment, with small values awarded for other things such as student surveys, attendance, and school encouragement for involving students and parents in education. The first final grades issued included 39 schools receiving an A, 198 receiving a B, 275 receiving a C, 250 receiving a D, and 69 receiving an F. Compared to preliminary FY11 school grades, 44 percent of school grades decreased in FY12.

Performance measures for public school support provide a snapshot of student performance generally when data is available after the end of the school year. Little or no consistent data is available through the year on student achievement and performance for state policymakers. For FY13, the Legislature appropriated $2.5 million for short cycle assessment for fourth through tenth grade students. To be meaningful, implementation should consider mandatory reporting to the Public Education Department at least three times a year, allowing policymakers access to data more than once annually. Additional benefits to intermediate reporting of student academic performance include (1) providing teachers the data necessary to alter instructional practices throughout the year to address student needs and (2) assisting the department in determining how to better support schools.

Research clearly demonstrates the importance teachers have on student learning. Despite a "highly qualified" teacher work force, improvement in student achievement is progressing slowly. The executive has proposed reforming the state's teacher evaluation system to measure the effect teachers have on student learning as measured primarily by student growth. Since 2010, the department has indicated the changes proposed require legislation; however, the federal government granted the state a waiver from certain federal No Child Left Behind provisions in exchange for implementation of an overhauled teacher and school leader evaluation system. To assist in implementation of a new evaluation system, the Legislature allocated $1 million to the PED for a new evaluation system based on student achievement growth. The PED promulgated regulations for a new evaluation system based on the following: 50 percent on student growth; 25 percent based on multiple observations; and 25 percent based on multiple measures. Data should be collected from public schools annually to allow districts and policymakers to address and improve school personnel policies concerning professional development, promotion, compensation, and tenure.

Measure	FY10 Actual	FY11 Actual	FY12 Target	FY12 Actual	Rating
Percent of fourth-grade students who achieve proficiency or above on standards-based assessments in reading	51.4%	46.5%	78%	49.9%	Y
Percent of eighth-grade students who achieve proficiency or above on standards-based assessments in reading	60.5%	53.3%	76%	54.3%	R
Percent of fourth-grade students who achieve proficiency or above on standards-based assessments in mathematics	45.4%	44.4%	77%	44.0%	R
Percent of eighth-grade students who achieve proficiency or above on standards-based assessments in mathematics	39.2%	40.8%	74%	41.7%	Y
Percent of recent New Mexico high school graduates who take remedial courses in higher education at two-year and four-year schools	47.1%	46.2%	40%	n/a	
Current year's cohort graduation rate using four-year cumulative method	67.3%	63%	75%	63%	R
				Program Rating	R

APPENDIX C: Teacher Effectiveness Analysis

Methodology

1. Using the Public Education Department (PED) Teacher-Student roster files from 2012, 2011, and 2010, imported SBA scaled scores for reading and math from the PED SBA data files. For each file, imported three-years of SBA data. Also, for the 2012 file, imported teacher preparation institution data from the PED licensure files.
2. Sent lists of teachers by institution to each institution to verify completer status. Moved teachers not verified by each institution into the "Other" preparation institution category.
3. Selected teachers in NM who have been teaching 8 or fewer years with at least 10 full academic year (FAY) students with valid SBA scores (not APA).
4. Calculated the difference from SY10 to SY12 scaled scores in both reading and math and analyzed the correlations between these two variables (DIFFM12M10 and DIFFR12R10) and student demographic variables to determine which to include in the linear regression model. Based on these correlations, included FRL in the regression model.
5. Ran two linear regressions, one for reading and one for math. For each, used the SY12 scaled score as the outcome variable. Used the SY10 scaled score, SY11 scaled score, and FRL as predictor variables. Calculated a predicted value and an unstandardized residual value (MathRes1012 and RdgRes1012).
6. For each of the reading and math unstandardized residual values, eliminated outliers greater than three standard deviations from the mean.
7. Aggregated the mean math and reading residuals by teacher, identified duplicates, and sorted by teachers with ten or more students. In excel, calculated a cumulative residual: for elementary teachers, calculated the mean of the reading and math residuals; for middle school math teachers, used the math residual; for middle school reading teachers, used the reading residual.
8. Repeated steps five through seven using teacher roster files from 2011 and 2010.
9. Averaged the mean residuals by teacher from 2012, 2011, and 2010 to create a 2012 value-added score by teacher.

Opportunities for future methodological improvements include bio-data matching of class rosters, using graduation data to match teachers to preparation programs, replacing values for missing SBA scores to eliminate selection bias, using other tested subjects in the regression equation, and converting of scaled scores to standardized scores (z-scores).

Regarding student gains from 2011 to 2012, two methods were applied: calculating the scaled score differences between 2011 and 2012, and adjusting the 2011 scores using Kelly's equation to reduce the spurious negative correlation between gains and 2011 scores. While adjusting the prior year scores reduced the r-value for reading and math, the overall mean gains between institutions were nearly identical with both approaches.

Demographics

- Of the approximately 23 thousand K-12 teachers in New Mexico in SY12, 2,879 met the following criteria:
- The teacher had eight or fewer years of teaching experience;
- The teacher could be connected with at least ten students in fourth through eighth grades with at least two years of math and reading SBA (not Alternative Proficiency Assessment) scaled scores; and
- The students connected to that teacher attended the institution for the full academic year (FAY) in SY12.

Those 2,879 teachers completed their training for initial licensure at Central New Mexico Community College (CNM), the College of Santa Fe (CSF), Eastern New Mexico University (Eastern), New Mexico Highlands University (Highlands), New Mexico State University (NMSU), Santa Fe Community College (SFCC), San Juan College (SJC), the University of New Mexico (UNM), and Western New Mexico University (Western). Because of

small numbers, students prepared at Northern New Mexico College (NNMC), Clovis Community College (CCC), and San Juan College (SJC) are reported in the "Other" category.

Given CNM's relatively new alternative licensure program, completers have the least experience, an average of 1.3 years, and are earning the lowest average annual salaries at $33 thousand. While the state average of Hispanic teachers is 38 percent, 65 percent of the teachers Highlands prepares are Hispanic and 55 percent are Hispanic at NMSU.

Table 17. Teacher Demographics by Prep Institute

Prep Institute	Number of Teachers	Average Salary	Average Years Experience	Male	Female	Caucasian	Native American	Hispanic	Other Ethnicity
CNM	31	$33,087	1.3	26%	74%	77%	0%	13%	10%
CSF	134	$42,521	3.0	15%	85%	60%	2%	34%	3%
Eastern	236	$39,009	3.8	15%	85%	62%	0%	38%	0%
Highlands	187	$38,374	3.2	17%	83%	34%	1%	65%	1%
NMSU	457	$40,352	3.8	20%	80%	43%	0%	55%	1%
SFCC	49	$37,470	3.3	22%	78%	80%	2%	18%	0%
UNM	990	$40,370	2.4	18%	82%	63%	3%	33%	1%
Western	64	$38,148	3.8	22%	78%	41%	3%	55%	2%
Other	731	$40,507	2.8	21%	79%	65%	2%	29%	4%
Total	**2,879**	**$40,086**	**3.0**	**19%**	**81%**	**58%**	**2%**	**38%**	**2%**

Source: LFC Analysis of PED Data

Statewide, 1,897, or 66 percent, of the teachers in this sample teach at the elementary level, 493, or 17 percent, teach middle school math, and 489, or 17 percent, teach middle school language arts. Compared with these state averages, CNM is preparing a higher percentage of secondary teachers, 55 percent, while Highlands is preparing a higher percentage of elementary teachers, 79 percent.

Table 18. Teacher Assignments by Prep Institute

Prep Institute	Elementary	Middle School Math	Middle School Language Arts
CNM	45%	39%	16%
CSF	67%	13%	20%
Eastern	69%	17%	14%
Highlands	79%	9%	12%
NMSU	68%	14%	18%
SFCC	59%	29%	12%
UNM	65%	17%	18%
Western	64%	22%	14%
Other	62%	20%	18%
Total	**66%**	**17%**	**17%**

Source: LFC Analysis of PED Data

In SY12, these 2,789 teachers had valid SBA scores for 97,045 students in grades four through eight. The demographic make-up of these students is representative of the overall population of K-12 students in New Mexico.

Table 19. Student Sample Demographic Profile

Category	Number of Students	Percent of Total
Male	48,646	50%
Female	48,399	50%
Caucasian	23,443	24%
Native American	7,604	8%
Hispanic	62,349	64%
Other Ethnicity	3,649	4%
Special Education	10,925	11%
English Language Learner	14,944	15%
Free or Reduced-Price Lunch	68,343	70%
Total	**97,045**	

Source: LFC Analysis of PED Data

Similarly, the breakdown of students by grade level and subject area allows for statistically significant conclusions in all three areas: 36,913 students or 38 percent were in elementary grades; 32,826, or 34 percent were in middle school math; and 27,306, or 28 percent, were in middle school language arts.

Of these 2,879 teachers, 548 also had valid mean residual values in SY11 and SY10 to calculate a three-year value-added score.

APPENDIX D: Principal Preparation Analysis

Demographics

Of the approximately 600 principals in New Mexico in SY12, 174 met the following criteria:

- The principal had not served in an administrative role in 2007. This metric was used to select new principals because the PED data set did not include a years of experience field exclusive to principal experience;
- The principal could be connected to the same school for all three years used to calculate SY12 school grades;
- Lists of principals were sent to the college listed as the institution of preparation within PED's licensure file. Institutions were provided with the opportunity to confirm that the principal had completed preparation through the institution. Multiple verification lists were sent to institutions. A few lists remained unverified. In these cases, principals were included.
- If an institution reported that a principal had not completed preparation through the institution, the principal was listed within the "other" category.
- Principals with unknown preparation institutions were sent to HED for preparation verification. In a few cases, the HED record agreed with the PED record, though the college of education rejected the principal as a completer. In these cases, the principal was included within the college's sample.
- Principals with verified administrator preparation institutions were included for analysis. Principals without verified institutions were listed within the "other" category for analysis.
  Principals trained out of state were also classified as "other."

Table 20. Principal Sample Demographics

Prep Institute	N	Average Salary	Male	Female	Caucasian	Native American	Hispanic	Other Ethnicity
Eastern	8	$72,188	63%	37%	88%	0%	13%	0%
Highlands	16	$64,839	44%	56%	38%	0%	56%	6%
NMSU	23	$71,173	57%	43%	48%	4%	48%	0%
UNM	33	$70,272	27%	73%	58%	3%	36%	3%
Western	19	$69,552	37%	63%	63%	5%	32%	0%
Other	75	$69,647	29%	71%	61%	5%	29%	4%

Source: LFC Analysis

Table 21. School Levels of Sampled Principals

	Elementary Schools	Middle and High Schools
Eastern	5	3
Highlands	8	8
NMSU	7	16
UNM	12	21
Western	10	9

Source: LFC Analysis

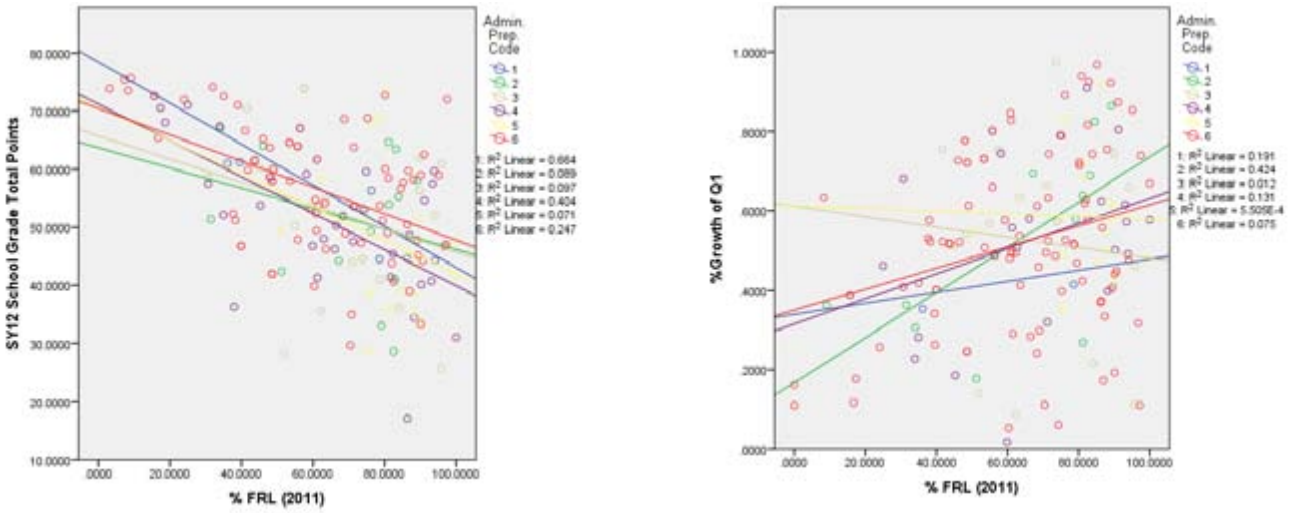
Methodology

Selected principals were matched to SY12 post-appeal school grades issued by PED. Principals connected to school grades 2.5 standard deviations above and below the mean were removed.

The 2012 school grades issued by NM PED are heavily influenced by poverty, with high poverty levels associated with low school grade totals. The percentage of students receiving free or reduced lunch (FRL) serves as a measure

of school poverty level. A school's FRL level was negatively associated several subcategories within school grades, including SY12 current status and Q3 growth, which measures the academic growth of the top 75 percent of students in a school. However, there was a slight positive correlation between FRL level and Q1 growth, which measures the growth of the lowest 25 percent of students in a school (r= 0.28, *p< 0.001*); as FRL level increases, so does the growth of a school's lowest performing students. Overall, a moderate negative correlation was found between a school's FRL level and 2012 school grade total (r= -0.50, *p< 0.001*); as school poverty increases, school grade totals decrease. Post-appeal school grades were used in analysis.

Chart 27. The Relationship between Poverty and School Grade Total and Growth of a School's Lowest Performing Students (Q1)



Source: LFC Analysis

A one-way ANOVA reveals a statistically significant difference in average SY12 total school grade-points attributed to New Mexico's institution Administration preparation programs ($F(5, 168)= 2.84$, $p= 0.017$)

Before controlling for poverty, differences among school grade subcategories also emerge when school poverty levels are not taken into account. A one-way ANOVA revealed significant differences among administrator preparation institutions within the school grade current status category ((F (5, 168) = 2.68, $p=0.026$) and growth of students within the top three quartiles ((F (5,168) =2.87, $p=0.016$)). No significant differences were noted among administrator preparation institutions within the school grade category that measures the growth of students in a school's lowest quartile.

After controlling for school poverty levels, however, school grade differences among programs appear much smaller. After adjusting for institution FRL levels using ANCOVA, there is no statistically significant difference in SY12 school grade totals among administrator preparation institutions.

ANCOVA

Dependent Variable: SY12 School Grade Total Points

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6761.868[a]	6	1126.978	10.460	.000
Intercept	74265.572	1	74265.572	689.312	.000
FRL2011	5091.870	1	5091.870	47.261	.000
Admin.Prep.Code	653.893	5	130.779	1.214	.305
Error	17776.889	165	107.739		
Total	517013.166	172			
Corrected Total	24538.756	171			

a. R Squared = .276 (Adjusted R Squared = .249)

ANCOVA

Dependent Variable: Growth Q3%

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.807[a]	6	.135	3.283	.005
Intercept	4.927	1	4.927	120.231	.000
FRL2011	.333	1	.333	8.130	.005
Admin.Prep.Code	.354	5	.071	1.726	.132
Error	6.557	160	.041		
Total	42.446	167			
Corrected Total	7.365	166			

a. R Squared = .110 (Adjusted R Squared = .076)

ANCOVA

Dependent Variable: %Growth of Q1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.747[a]	6	.125	2.825	.012
Intercept	1.808	1	1.808	40.998	.000
FRL2011	.512	1	.512	11.618	.001
Admin.Prep.Code	.145	5	.029	.660	.654
Error	7.055	160	.044		
Total	53.605	167			
Corrected Total	7.803	166			

a. R Squared = .096 (Adjusted R Squared = .062)

When principal performance is disaggregated according to school level, significant differences among elementary principals emerge while the patterns observed among the principal population as a whole persist among secondary principals. A one-way ANOVA revealed no statistically significant differences in school grade-point totals or growth of the bottom quartile of students among institutions that prepared elementary school principals. However, a one-way ANOVA suggests that a statistically significant portion of the variance in growth of the top three quartiles of students may be explained by a principal's institution of preparation (F (5, 62) = 2.69, $p= 0.029$). While correlations between school poverty level and other measures within school grade totals persist, no statistically significant correlation between the growth of the top three quartiles of student and school poverty level exists, which suggests that there are meaningful differences in the growth of the top three quartiles of students that may be attributed to elementary principals from different institutions.

A Pearson correlation revealed no statistically significant correlation between principal annual salary and school grade, suggesting that principals with more experience are not connected with schools that earn higher grades within New Mexico's school grading system. (r= -0.33, $p= 0.664$). Other variables that were found to have no significant correlation with school grade measures include principal NMTA score (r=0.102, $p=0.376$) and whether

or not the principal led the school in the year prior to school grade data collection (r=0.004, $p=0.955$). Additionally, no significant correlations were found between principal salary, NMTA score, and school poverty level.

Principal experience is not related to school performance, as measured by NM PED school grades. A second principal sample which included all principals who were present at the same school site between SY10 and SY12 was similarly analyzed. This analysis revealed trends like those observed among the sample of principals which only included recently prepared principals; no statistically significant differences between programs were observed after the effects of poverty were controlled. Principal salary was used as a proxy for experience in this analysis, as principal salaries generally increase with years of experience.

Teacher Perception of Preparation. Teacher surveys were sent to every superintendent and director of human resources in the state with the request that questionnaires be distributed to all teachers. Surveys were also sent directly the email addresses provided by PED. The LFC received 4,079 teacher responses.

The majority of teachers prepared by New Mexico's publicly funded institutions report feeling adequately prepared to teach. Among programs, there are significant differences in the degree to which teachers feel prepared, particularly to teach reading and meet the needs of diverse students, but teachers generally agree that their program prepared them for classroom realities.

Table 22. Teachers Who Report Feeling "Well" or "Sufficiently Prepared" by Their Program of Preparation

	Manage the Classroom	Teach Reading	Teach Math	Support Students with Disabilities	Teach ELL Students	Use Student Data
CNM	88%	88%	71%	82%	72%	93%
Eastern	73%	71%	73%	62%	46%	61%
Highlands	80%	70%	70%	68%	69%	58%
NMSU	69%	62%	68%	55%	43%	51%
UNM	66%	61%	67%	56%	51%	56%
Western	78%	66%	63%	60%	51%	68%

Highlighted cells indicate statistically significant differences at the p=0.05 level.

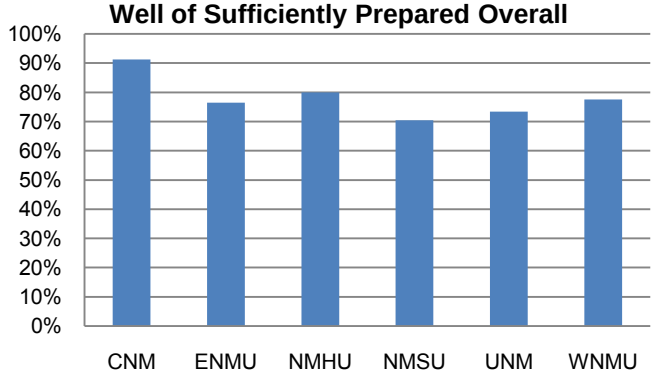
Source: LFC Survey

Principal Perceptions of Teacher Preparation. The state's principals were surveyed to collect their perceptions of the quality of teacher candidates produced by the state's preparation programs and their own administrator preparation programs. Attempting to reach every administrator in New Mexico, surveys were sent directly to the emails of 640 principals. Of these, 213 principals responded.

Teachers and principals tend to disagree about the programs that best prepare teachers. While surveyed teachers prepared by alternative licensure programs (including CNM, NNMC, SFC, and SJC) report feeling more prepared, principals consistently report that alternative licensure candidates are less prepared than completers of New Mexico's five traditional teacher preparation programs. Additionally, while teachers from UNM report feeling less prepared than completers of other programs in the state, principals report that UNM prepares the highest quality candidates.

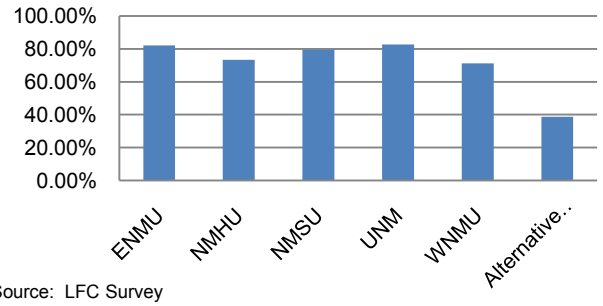
The majority of survey respondents agree that teachers from all traditional preparation programs are prepared overall. Agreement levels were highest for teachers prepared by Eastern (82 percent) and UNM (83 percent). Only 39 percent of principals rate alternative licensure completers as prepared overall.

Chart 28. Teachers Who Report Feeling Well of Sufficiently Prepared Overall



Source: LFC Survey

Chart 29. Principals Who Agree "Teachers Prepared by New Mexico's Schools of Education are Well or Sufficiently Prepared Overall"



Source: LFC Survey

When asked about the importance of various elements of teacher preparation, principals indicate that classroom management and student teaching are most critical.

Table 23. Principals Who "Strongly Agree" that Teacher Preparation Experiences Are Critical

Early field experiences	64%
Lengthy field experiences	54%
Student teaching	80%
Content knowledge	75%
Pedagogical knowledge	59%
Knowledge related to data collection and analysis	60%
Knowledge in meeting the needs of diverse learners	64%
Strategies for classroom management	86%
Cultural awareness and strategies for appropriate interaction	53%

Source: LFC Survey

Principals generally rate themselves as well prepared, though they report being less prepared for specific competencies, including using data effectively and designing professional development.

Table 24. Principals Who Agree They Were Well or Sufficiently Prepared for Administrator Duties

Competencies	Eastern	Highlands	NMSU	UNM	Western
Evaluate Curriculum	85%	87%	62%	89%	74%
Evaluate Teachers	85%	82%	72%	85%	74%
Use Data to Monitor Progress	62%	87%	38%	72%	74%
Design Professional Development	67%	83%	59%	76%	74%
Manage School Operations	92%	91%	90%	89%	79%
Engage the Community	92%	83%	76%	83%	80%
Make Ethical Decisions	100%	100%	100%	98%	100%
Respond to Community Context	100%	83%	93%	91%	84%
Serve as Instructional Leader	92%	87%	93%	96%	90%
Establish Institution Culture	92%	91%	93%	96%	90%
Prepared Overall	**92%**	**87%**	**93%**	**93%**	**84%**

Highlighted cells indicate statistically significant differences at the p=0.05 level.

Source: LFC Survey

Superintendent and Directors of Human Resources. Surveys were sent to all superintendents and directors of human resources in the state. After filtering the 93 responses to remove survey completers who were not involved in the hiring of principals, 18 survey responses remained.

Survey results reveal significant differences in district administrator perceptions of principal preparation among New Mexico's colleges of education. A one-way ANOVA was used to test differences in perceptions of principal preparation among colleges of education. Perception of overall principal preparation differed significantly across colleges of education the (F (4, 47) = 2.70, $p = 0.043$). District-level administrators with experience hiring principals report that Eastern and UNM produce principals who are most prepared overall, while Highlands and Western produce principals who are least prepared overall.

Table 25. District Administrators Who Agree "Principals are Well or Sufficiently Prepared Overall"

University	Mean Response (1-4 Scale)
Eastern	3.08
Highlands	2.33
NMSU	2.75
UNM	3.08
Western	2.5

Source: LFC Survey

No statistically significant differences among programs emerged when administrators were asked about elements of preparation, including evaluating teachers and curriculum, developing a positive institution Culture, managing school operations, and serving as an educational leader.

APPENDIX F: Clinical Experience Rubric

	Minimum Standard	Exemplary Standard
Entrance Requirements	Candidates have completed specified coursework as determined by the institution	Placement is not automatic. Candidate demonstrates competence in the following areas prior to student teaching: lesson plan development, summative and formative assessment development, analyzing student data, effective reading instruction, a small repertoire of classroom management skills
Timing	A portion occurs within the first 30 credit hours Student teaching takes place within the senior year, when a candidate assumes responsibility for a class	Clinical experiences are integrated throughout the preparation to allow candidates to apply theory as it is learned
Placement Procedures	Collaborative relationships exist between colleges of education and placement sites, leading to a sense of shared responsibility and accountability College of Education plays a role in supervising teacher selection and approval	Student teachers or interns have the opportunity to develop skills in more than one school level and demographic setting. Specifically, student teachers should have opportunities to experience placement in high-performing, high-poverty schools
Supervision	Student teachers are under the direct supervision of a teacher	Student teachers are supervised by both university faculty and rigorously selected and prepared. Mentor teachers have a minimum of three years of teaching experience, have demonstrated their effectiveness via measures of student achievement, and have either undergone training in effective mentoring or have demonstrated their effectiveness as mentors.
Observation	Student teachers are observed and have the opportunity to observe others.	Student teachers are observed a minimum of five times during their student teaching experiences by both university faculty and mentor teachers.
Opportunities for Feedback	Student teachers and interns are provided with feedback	Student teachers are provided with a conference and written feedback from university faculty and mentor teacher after every observation
Length	14 weeks	one year, full time
Assessment	Candidates are provided with experiences to reflect upon their own knowledge and skills. They complete summative and formative assessments that demonstrate mastery of New Mexico's new teacher competencies. Students demonstrate mastery of beginning teacher competencies through a capstone project that includes a portfolio/ action research project that demonstrates a teacher's ability to analyze student data and alter instructional strategies to improve student outcomes.	Candidates are evaluated according to student achievement and student data, including student artifacts, summative and formative assessments

Materials reviewed: EARS, field experience manuals and syllabi submitted by institutions
1= Meets minimum standard
2= Somewhere in between
3= Meets exemplary practice standards

Sources:

Boyd, Donald J., Pamela L. Grossman, Hamilton Lankford, Susan Loeb, and James Wyckoff. "Teacher Preparation and Student Achievement." *Educational Evaluation and Policy Analysis* 31.4 (2009): 416-440.

Levine, Arthur. "Educating School Teachers." *The Education Schools Project* (2006): 1-142.

National Council For Accreditation Of Teacher Education. "Transforming Teacher Education Through Clinical Practice: a National Strategy to Prepare Effective Teachers." (2010): 1-30.

"Student Teaching in the United States: Key Ingredients for Strong Student Teaching." National Center For Teacher Quality, 2011. Web. July 2012. <www.nctq.org/edschoolreports/studentteaching>

Section 22-10A-6 NMSA 1978

NCATE Accreditation standards

Standard 1: **Candidate Knowledge, Skills, and Professional Dispositions**	Candidates know and demonstrate the content knowledge, pedagogical content knowledge and skills, and professional dispositions necessary to help all students learn. Assessments indicate that candidates meet professional, state, and institutional standards as 80 percent or more of the program's completers pass the state's licensing examination.
Standard 2: **Assessment System and Unit Evaluation**	The preparation program has an assessment system that collects and analyzes data on applicant qualifications, candidate and graduate performance, and program operations to evaluate and improve the performance of candidates, the institution, and its programs.
Standard 3: **Field Experiences and Clinical Practice**	The program and its school partners design, implement, and evaluate field experiences and clinical practice so that teacher candidates and other school professionals develop and demonstrate the knowledge, skills, and professional dispositions necessary to help all students learn.
Standard 4: **Diversity**	The program designs, implements, and evaluates curriculum and provides experiences for candidates to acquire and demonstrate the knowledge, skills, and professional dispositions necessary to help all students learn. Assessments indicate that candidates can demonstrate and apply proficiencies related to diversity. Experiences provided for candidates include working with diverse populations, including higher education and P–12 institution Faculty, candidates, and students in P–12 schools.
Standard 5: **Faculty Qualifications, Performance, and Development**	Faculty are qualified and model best professional practices in scholarship, service, and teaching, including the assessment of their own effectiveness as related to candidate performance. They also collaborate with colleagues in the disciplines and schools. The program systematically evaluates faculty performance and facilitates professional development.
Standard 6: **Unit Governance and Resources**	The program has the leadership, authority, budget, personnel, facilities, and resources, including information technology resources, for the preparation of candidates to meet professional, state, and institutional standards.

Report
to
The LEGISLATIVE FINANCE COMMITTEE

Public Education Department
Promoting Effective Teaching in New Mexico
November 15, 2012

Report #12-12

LEGISLATIVE FINANCE COMMITTEE

Senator John Arthur Smith, Chairman
Representative Luciano "Lucky" Varela, Vice-Chairman
Senator Sue Wilson Beffort
Senator Pete Campos
Senator Carlos R. Cisneros
Representative William "Bill" J. Gray
Senator Stuart Ingle
Representative Rhonda S. King
Representative Larry A. Larrañaga
Senator Carroll H. Leavell
Senator Mary Kay Papen
Representative Henry "Kiki" Saavedra
Representative Nick L. Salazar
Representative Edward C. Sandoval
Senator John Sapien
Representative Don L. Tripp
Representative James P. White

DIRECTOR

David Abbey

DEPUTY DIRECTOR FOR PROGRAM EVALUATION

Charles Sallee

PROGRAM EVALUATION TEAM

Jeff Canney, CGFM
Jon R. Courtney, Ph.D.
Valerie Crespin-Trujillo
Jack Evans
Brenda Fresquez, CICA
Pamela Galbraith
Maria D. Griego
Rachel Mercer-Smith
Matthew Pahl
Michael Weinberg, Ed.D.

State of New Mexico
LEGISLATIVE FINANCE COMMITTEE
325 Don Gaspar, Suite 101 • Santa Fe, NM 87501
Phone: (505) 986-4550 • Fax (505) 986-4545

David Abbey
Director

November 15, 2012

Ms. Hanna Skandera, Secretary-Designate
Public Education Department
Jerry Apodaca Education Building
300 Don Gaspar Avenue
Santa Fe, NM 87501

Dear Ms. Skandera:

On behalf of the Legislative Finance Committee (Committee), I am pleased to transmit the evaluation, *Effective Use of Student Test Data to Assess & Improve Teacher Evaluation.* The program evaluation team assessed the three-tiered licensure system; value-added models; and resource allocation from the state funding formula. The report will be presented to the Committee on November 15th, 2012. Exit conferences were conducted with the Public Education Department on November 7th, 2012 to discuss the contents of the report. The Committee would like a plan to address the recommendations within this report within 60 days from the date of the hearing.

I believe this report addresses issues the Committee asked us to review and hope New Mexico's public education system benefits from our efforts. We very much appreciate the cooperation and assistance we received from your staff.

Sincerely,

David Abbey, Director

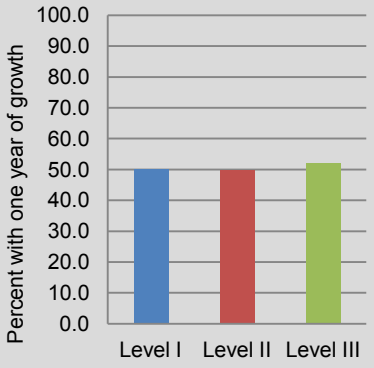
Cc: Senator John Arthur Smith, Chairman, Legislative Finance Committee
Representative Luciano "Lucky" Varela, Vice-Chairman, Legislative Finance Committee
Representative Henry "Kiki" Saavedra, Legislative Finance Committee
Representative Rick Miera, Chairman, Legislative Education Study Committee
Ms. Yolanda Berumen-Deines, Secretary, Children, Youth, and Families Department

Table of Contents

Page No.

Students in high-performing teachers' classrooms gain far more than a year's worth of academic growth, increasing an average of three points on the reading SBA and five points on the math SBA.

Percent of Students With One Year of Growth (SBA Math Gain Scores)



Source: LFC Analysis of PED Data

Nearly 30 percent of level III teachers, the state's highest paid, are in the bottom quartile of teacher performance in math and reading.

Quality teaching is the most influential school factor affecting academic success. States and school districts across the country increasingly recognize this and create incentives to improve teaching quality. One such effort involves using student performance information, through value-added models (VAM), to evaluate teaching effectiveness.

In 2003, New Mexico introduced the three-tiered system to increase the recruitment and retention of quality teachers to improve student achievement. The system created a three-level career ladder for teachers to ascend based on experience, leadership, and skills. Movement up a level results in pay increases of $10 thousand. Previous evaluations of the three-tiered system confirmed the system decreasing widespread teacher shortages, reducing unqualified teachers, and improving teacher pay.

Student performance, however, has not improved with taxpayer investments in teacher pay. A 2009 Legislative Finance Committee (LFC) evaluation using one year of performance data confirmed small differences in performance despite large differences in pay among teachers and offered solutions for improvement. The recommendations were not implemented. Since that time, nearly 6,000 teachers advanced to new license levels, receiving $59 million in mandatory salary increases.

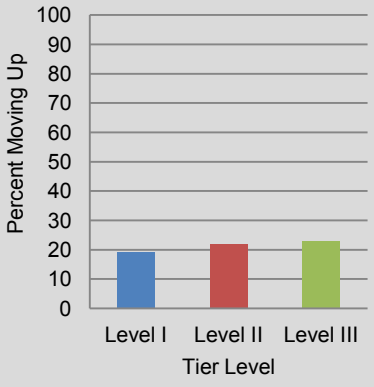
This evaluation assessed the status of the system since the majority of teachers have advanced at least one level and reviewed options for using VAMs to identify effective teaching. The evaluation used multiple years of student and teacher data to assess the performance of New Mexico's fourth through eighth grade teachers and partnered with researchers at the University of New Mexico to model how student populations influence VAM calculations.

Student performance within teacher licensure levels and between licensure levels suggests the local and state evaluation systems are not screening teachers for their effectiveness in the classroom. The difference in performance between teachers of each of the three levels is small. For example, 50 percent of students taught by level I teachers achieved a year's worth of growth in math in 2012, compared to 52 percent of students of level III teachers. Furthermore, each licensure level has high and low performing teachers; in 2012, nearly 30 percent of the lowest performing reading and math teachers in the state had a level III license. These teachers can maintain their level, including those grandfathered into the system, for the rest of their careers because the local evaluation and state license renewal process lacks factors for student achievement.

The three-tiered system continues to offer a solid framework to align resources to performance, but student achievement must be better incorporated into the process. If modified, student achievement could be a data-driven concern for all teachers and serve as a way to reward the state's best teachers and intervene for struggling teachers. While lack of

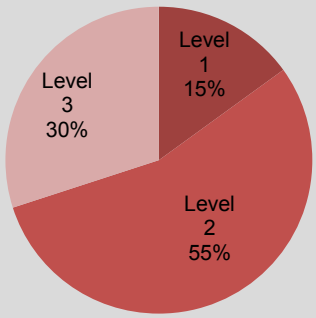
Unlike in 2003, when standardized testing was relatively new, the state now has many years worth of student and teacher data to be incorporated into the three-tiered system and funding formula.

Percent of Students Moving Up At Least One Proficiency Level in SBA Reading Scores, SY12



Source: LFC Analysis of PED Data

License Levels as a Proportion of Low Performing Reading Teachers SY12



Source: LFC Analysis of PED Data

longitudinal data made it difficult to use student performance in teacher evaluations when the three-tiered system was implemented in 2003, the state now has the resources and expertise to incorporate that information. Properly implemented, VAMs can identify teachers for advancement; their complexity, however, limits VAMs role in annual local evaluation of teachers.

PED has sought, through rule, to improve the local evaluation component of the three-tiered system and initiated two task forces to examine how to incorporate student achievement, including using VAMs, into a new system. However, statutory changes not addressed by the new PED rule are necessary to reform local evaluations and the state licensure system.

Finally, the state has not incorporated the three-tiered system into the funding formula. Instead, the formula uses a district-wide training and experience (T&E) factor, even though districts with high T&E values do not regularly achieve better performance than those with low T&E values. As currently structured, T&E widens the achievement gap by providing more funding for more affluent school districts.

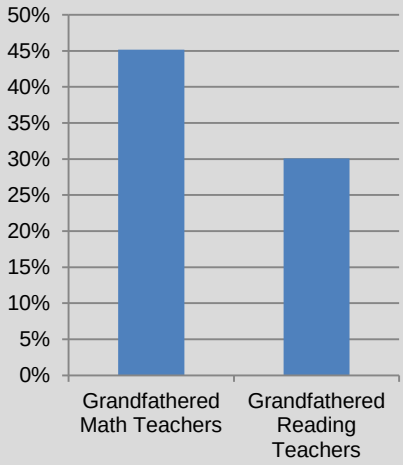
KEY FINDINGS

New Mexico's three-tiered career ladder system does not align pay with student achievement. Student performance within teacher licensure levels and between licensure levels suggests local and state evaluation systems are not screening teachers for effectiveness in the classroom. The difference in performance between teachers of each of the three licensure levels is small, with many high and low-performing teachers at each level. Teachers maintain levels throughout their careers because student achievement is not factored into licensure renewal. Establishing expectations for student achievement in the local and state evaluation systems will better align pay with student achievement.

Improving student achievement was a key policy goal of implementing the three-tiered system. The three-tiered system's founding legislation identifies student success as the fundamental goal of New Mexico's education system. The three-tiered system was designed to help achieve this goal by attracting, retaining, and holding accountable quality teachers.

The state has not established expectations for student achievement in evaluation of level I, II, and III teachers. Competencies used in the state and local evaluations of the three-tiered system include examples of student performance, but the evaluations have no expectations for the performance of all students, particularly on standardized tests. When the three-tiered system was established, the SBA was new and lacked longitudinal information; student performance, therefore, was not incorporated into evaluations. Teachers at different license levels achieve similar student performance, and a majority of New Mexico teachers do not feel the state evaluation process identifies effective teachers.

Grandfathered Teachers with Less than Half of Students Obtaining a Year's Worth of Growth by Subject



Source: LFC Analysis of PED Data

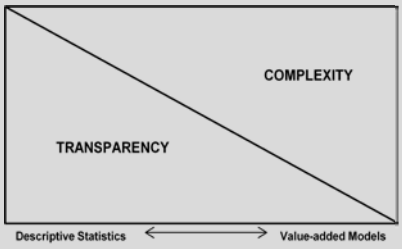
The local evaluation system does not differentiate between high and low-performing teachers or focus on student achievement. Evaluation requirements need strengthening to align with the common core, use student data, assess teacher effectiveness, and improve satisfaction among teachers. PED has not revised the competencies and requirements used in the state's professional development dossier and local evaluation system since implementation and the process could be updated to better reflect current education research.

The professional development dossier (PDD) does not effectively screen teachers for advancement, resulting in ineffective teachers receiving large pay increases. As a result, high and low performing teachers exist at each licensure level. The lack of clear and consistent performance among teachers in each licensure level shows the PDD process does not reward a teacher's impact on student achievement.

The state allows low-performing teachers to keep their license level because the state does not have a rigorous license renewal process. Grandfathered teachers, those that obtained a level II or III licensure without going through the state's PDD, continue to renew their licenses without passing the PDD. Many of these are low-performing teachers with the highest mandated salary in the state.

The three-tiered system offers a framework to align resource allocation to performance, but student achievement must be better incorporated into the process. The system offers significant salary increases and a competency structure; if modified, this system could drive student achievement across the state by setting student performance expectations. A modified three-tiered system could strategically reward the state's best teachers and provide strategic interventions for struggling teachers.

Understandability of Statistical Models



<u>When used appropriately, value-added models (VAMs) can help identify teachers' success levels and drive student achievement.</u> Different VAMs can show volatility among certain teachers. This can be minimized by using two different VAMs to form a composite score. Once these are controlled for, VAMs can be responsibly used to reward outstanding teachers and help those who are ineffective.

Value-added models are increasingly used across the country to evaluate teacher performance. As of October 2012, the U.S. Department of Education granted 33 states, including New Mexico, a waiver from some of the requirements of No Child Left Behind for changing their teacher evaluation systems to incorporate student data. Many proposals included use of VAMs for 50 percent of a teachers' evaluation rating. VAMs have the potential to inform stakeholders about teacher performance, but the volatility in these models warrants caution moving forward because of potential misclassification of teachers.

Depending on the demographic factors used, value-added models produce varied results. Some VAMs control for demographic factors and use multiple years of scores on a handful of different assessments, while others

How Different Value-added Models Using One-year of Data Affect a Teacher with a High Proportion of At-Risk Students

Test Score Only Model (no student demographic factors)	
Math	**Reading**
Needs Improvement	Needs Improvement
Student Demographic Model (includes all available student demographic factors)	
Math	**Reading**
Highly Effective	Highly Effective

Source: UNM

Value-added models effectively identify very-high and very-low performing teachers.

do not. Experts continue to debate about the usefulness of these different models. In 2012, Pearson Education, Inc. published a study comparing five different VAM teacher evaluation approaches and concluded that the results are not definitive and depend on the model used.

Some value-added models adversely affect educators teaching certain populations of students. The Value-added Model Research Group at the University of New Mexico's College of Education used five years of teacher and student data to determine scores for teachers from two different VAMs on teacher scores. One VAM incorporated only test scores (test-score only model), while the other compensated for contextual variables, such as poverty and English language fluency (student demographic model).

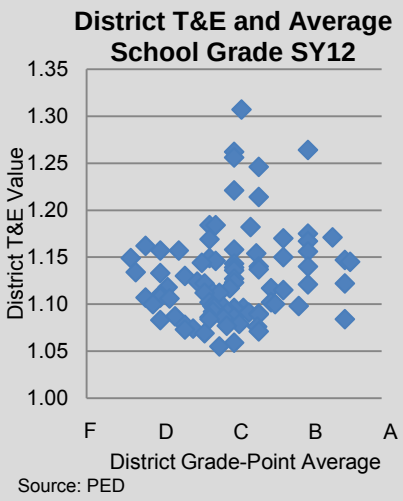
Value-added models are limited in what they can tell educators, the public, and other stakeholders. VAMs compare teachers with one another, making them a relative concept. One set of scores from VAMs do not indicate whether quality instruction is occurring in classrooms, only whether one teacher has students with higher achievement gains than another. Over time, VAM scores across multiple years can create a picture of absolute teacher performance that can be used for licensure advancement.

The use of value-added models can be responsibly integrated into the three-tiered system to identify teachers for advancement and bonus pay. VAM methodologies being developed for teacher evaluation and school grading could be leveraged to measure performance at tier levels and used as benchmarks in advancement between levels. Specifically, VAMs can be used to reward good teachers and identify poor teachers for professional development as part of a larger system of teacher evaluation.

Resource allocation amongst districts and schools creates funding disparities between low-income students and their more affluent peers, without driving student achievement. The funding formula rewards more affluent districts through the T&E index, a multiplier that allocates funds toward districts based on two variables that have not been shown to affect student achievement: teacher education levels and experience. Aligning this factor with a modified three-tiered system and offering a mix of incentives for high-performing teachers in low-income schools will better align resources with student achievement.

New Mexico directs nearly $200 million for higher teacher compensation through the T&E index in the public school funding formula. T&E index values are based on teachers' highest degree earned and years of experience. That index value is multiplied by student demographic and program units already generated in the formula. Funds generated by the T&E index in the funding formula compose up to 20 percent of a district's formula funding.

The T&E index directs more funding to more affluent school districts and produces a questionable return on investment after factoring in poverty. The T&E does not recognize better performance by teachers and higher pay, but instead rewards relatively affluent districts for keeping teachers and sometimes requiring them to meet higher education requirements. Aligning



Source: PED

Average Salaries at High and Low-Poverty Elementary Schools in Selected Districts SY12

Dist.	Elementary School	Avg. Salary	FRL
APS	Emerson	$42,900	96%
	Georgia O'Keefe	$46,923	15%
SFPS	R.M. Sweeney	$41,503	87%
	Wood-Gormley	$49,506	19%
LCPS	Booker T. Washington	$43,714	90%
	White Sands	$46,385	30%

Source: PED

the T&E index to a modified three-tiered system that focuses on student performance will allow the state to send resources to high-performing teachers and schools. High-poverty, rural districts with the greatest needs generally have difficulty hiring experienced teachers with advanced degrees who increase the T&E index. As FRL levels increase, T&E values drop.

The T&E index is not aligned to the three-tiered system. The three-tiered system provides large salary changes not accounted for in the T&E index. For several years the LFC has noted, in its present form, the T&E index is not aligned to the three-tiered system. An evaluation of the public school funding formula conducted jointly by the LFC and the Legislative Education Study Committee, the New Mexico Effective Teaching Task Force final report, and the AIR funding formula study recommended better alignment of the T&E index with the three-tiered system.

Level III teachers are more likely to teach in more affluent districts and schools. While districts do not have explicit policies to move more experienced and educated teachers to more affluent schools, students in poverty are more likely to have a less experienced, poorer performing teacher. Recruiting and retaining high quality teachers in low-income schools is integral to ensuring students in poverty achieve academically.

The state system can offer a mix of incentives to recruit and retain good teachers in high-poverty schools. Research shows financial incentives can recruit high-quality teachers and slightly decrease turnover in the short-term, but money does not work in the long-term to keep teachers at low-income schools: "Even when bonuses succeeded in drawing teachers to the poorest schools, such incentives could not compensate for the lack of support they encountered in these schools, which in turn contributed to the departure of many of these teachers."

KEY RECOMMENDATIONS

The Legislature should:

Replace the PDD and establish the effective teaching portfolio (ETP) as part of the licensure advancement application with new requirements and competencies. The ETP will have new requirements strengthening expectations for student achievement, requiring satisfactory annual evaluations, and allowing the most effective teachers, as measured by a statewide VAM, to bypass the ETP process;

Create licensure terms for level I, II, and III licenses. Level I licenses should have a five-year license term; and teachers must submit for renewal after three years; level II and Level III licenses should have an eight-year term and teachers must submit for renewal after six years;

Create new requirements for level II and level III licensure renewal, including meeting student performance expectations through the ETP or statewide VAM, and allowing teachers not meeting those expectations extra time to show competency before being denied renewal of a teaching license;

Require PED to annually rank the performance of licensed teachers providing instruction in tested grades and subjects through two different value-added models;

Change the T&E index to an effective teacher index that rewards districts based on the number of teachers they have in each license level;

Consider a mechanism, possibly through the funding formula, to provide additional compensation to effective teachers (as measured by the new aforementioned teacher evaluation and three-tiered licensure system) to teach in high-poverty schools.

PED should:

Establish updated basic competency and effectiveness indicators for teachers, as part of the ETP, including setting new student achievement expectations for Level II teachers and more rigorous student achievement expectations Level III teachers;

Establish an effectiveness evaluation for career teachers (level II and III), occurring every three years that incorporates student achievement and professional development goals. Public schools may use the results of the effectiveness evaluation to make employment decisions, in accordance with other provisions of law;

Require a professional development plan by the 40th day establishing the current year's performance goals, including measurable objectives for student performance. The goals should be based on updated basic competency and effectiveness indicators, the previous year's annual evaluation, and a previous year's students' performance;

Create and use a statewide VAM that uses two different calculations to obtain a composite score to help eliminate VAM biases for teachers of certain populations;

Provide that a performance evaluation be conducted annually for all teachers, and be based on whether a teacher exceeds, meets or, does not meet expectations on basic competency and effectiveness indicators, professional development goals, and satisfaction from parents. All teachers must be observed by principals 3 times a year.

Overview of the three-tiered system. In 2003, the Legislature passed comprehensive education reform, including the establishment of the three-tiered system and corresponding new minimum salaries. The School Personnel Act of the Public School Code outlines the three-tiered system certification and compensation schedules. The minimum salaries established in law were phased in between 2003 and 2008:

- Level I, Provisional Teacher: $30,000 in SY04;
- Level II, Professional Teacher: $35,000 in SY05 and $40,000 in SY06; and
- Level III-A, Master Teacher: $45,000 in SY07 and $50,000 in SY08.

The three-tiered system requires teachers to submit a professional development dossier (PDD) for level advancement. The PDD is a collection of evidence of teacher performance assessed by external reviewers, and is intended to provide sufficient evidence that a teacher is qualified to advance to a higher licensure level. The PDD evaluates teachers on nine teacher competencies.

Fast Facts of the three-tiered system and the PDD process

The PDD process has three submittal periods per year: February, June, and November, and takes about three to four months to complete

Each PDD submittal costs a teacher approximately $100.

2012 PDD pass rate:

- Teachers moving from level I to level II: 92 percent
- Teachers moving from level II to level III: 90 percent

Not all teachers at higher licensure levels submitted a PDD for advancement, as over 2,700 teachers advanced to Level III between the effective date of HB 212 in April 2003 and the effective date of rules requiring teachers to submit PDD's for advancement in July 2004. This gap in timing provided a window for certain teachers to qualify for $50 thousand salaries without submitting a PDD. Many of these "grandfathered" teachers are still actively teaching in New Mexico public schools and hold a level II or III license without submitting a PDD.

The three-tiered system has substantially increased teacher pay in New Mexico. According to a report issued by the National Education Association (NEA), salaries for New Mexico public school teachers increased 38.8 percent from the SY01 to SY11. This is the eighth highest increase among states in the nation during the ten year time span. In SY11, New Mexico ranked 40th among states, with an average public school teacher salary of $46,888, according to the NEA report. The report did not factor in cost of living into its rankings.

Evaluations as part of the three-tiered system. Local and state evaluations are both components of the three-tiered system. School administrators conduct local evaluations every one to three years and focus on evaluating classroom practice. State evaluations are conducted through the PDD, and teachers submit local evaluations as a part of the PDD when applying to advance a license level. While the local and state evaluations use the same competencies to evaluate teachers, only local evaluations are conducted at school sites.

Local evaluations are also considered for teachers renewing their licensure. Based on local annual evaluations, the superintendent of the district (or governing authority of other institution or school) verify the applicant has demonstrated the competencies for the current level of licensure and has met other requirements of the state's highly objective uniform standard of evaluation (HOUSE).

Three-tiered system and improved student performance. One of the primary purposes of the three-tiered system was to facilitate student success by recruiting and retaining qualified teachers. In the 2003 *Assessment and Accountability Act*, the legislative findings and purposes section begins and ends with an emphasis on student success. The legislation states, "The key to student success in New Mexico is to have a multicultural education system that attracts and retains quality and diverse teachers to teach."

Given one of the primary purposes of the three-tiered system is to ensure student success, it is appropriate to explore the connections between advanced licensure levels and increases in student performance. The standards-based assessment (SBA) is a statewide assessment given annually to third through eighth-grade students and again to eleventh-graders. The SBA meets the requirements of No Child Left Behind and is based on New Mexico state standards. New Mexico has four levels of performance used by the SBA: beginning step, nearing proficiency, proficient, and advanced. Proficient performance is expected of New Mexico students.

Other reports have explored the connection between the three-tiered system and student performance, including a Legislative Education Study Committee (LESC) memo in 2006 describing the extent to which the three-tiered system requires documentation of student achievement. In 2007, a joint evaluation by the Office of Educational Accountability (OEA), the LESC, and the LFC suggested further study into the links between advanced licensure and student academic performance. A 2009 joint report by the Legislative Finance Committee, Legislative Education Study Committee, and the Office of Education Accountability studied links between licensure and student performance. This report serves as a follow-up to the 2009 report.

Data used for LFC three-tiered analysis

The LFC used teacher records linked to student math and reading SBA scores to analyze the relationship between licensure level and student achievement. The total numbers of 3-8 grade teachers with at least one math or reading score ranged between 6,900 and 7,200 teachers. Teachers with fewer than 10 students were eliminated from the dataset along with level 0 teachers and students who did attend the same school for the full academic year (FAY). Non-FAY students were included in subsequent VAM analysis. Teachers with only third grade students were eliminated from the dataset as the students for these teachers did not have prior-year data to calculate gain. Data below reflect the populations after these data cleaning techniques were applied.

Selected Descriptive Statistics for New Mexico Teachers Used in LFC three-tiered Analysis

Year	Number of Teachers	Average Salary	Average Years of Experience	Percent Hispanic	Age in years at start of year
SY10	4,629	$45,612	9.7 years	31.3%	44.3
SY11	4,608	$45,531	10.1 years	36 %	44.1
SY12	4,595	$44,788	10.1 years	35.8%	44.3

Source: LFC Analysis of PED Data

Selected Descriptive Statistics for New Mexico Teachers Used in LFC three-tiered Analysis

Year	Number of Students	Percent FRL	Percent ELL	Percent Hispanic	Percent SPED	Percent Gifted
SY10	98,378	68.5%	13.7%	57.7%	11.6%	7.9%
SY11	101,029	69.5%	13.7%	61.1%	11.3%	8%
SY12	102,152	69.8%	14.6%	61.5%	11.7%	8.7%

Source: LFC Analysis of PED Data

The three-tiered system and the training and experience (T&E) multiplier has a major impact on a district's allocation from the State Equalization Guarantee (SEG). This multiplier increases funding allocated to a district by as much as 20 percent based on teaching staff credentials and experience. High-poverty, rural districts with the greatest needs generally have the greatest difficulty hiring experienced teachers with advanced degrees and receive less funding from the T&E index. For several years the LFC has noted that the T&E index is not aligned to the three-tiered system.

Evaluation and value added modeling. States are increasingly relying on a statistical procedure known as value added modeling (VAM) to evaluate teacher performance. VAMs have the potential to inform stakeholders when student achievement data exists (reading and math) but not other subject areas. There is not one widely accepted VAM among education researchers or administrators. Value added models use data from students' past test scores to predict subsequent scores and then subtracts that prediction from current year scores to provide an estimate for teachers. This estimate is the "value added" and the models themselves can range from simple statistical procedures to more complex, multi-level models. Models can be run in basic statistical software, but more complex models require custom programming of statistical formulas and are increasingly being run by private, for-profit corporations.

Development and reliance on VAMs for teacher evaluation is controversial. Teachers in Chicago Public Schools went on strike in part because of reliance of the city on VAMs accounting for 45 percent of teacher evaluations. After seven lost school days, a renegotiated contract reflected a lesser reliance on VAMs, accounting for 30 percent of teacher evaluations.

Formed by Executive Oder in April 2011, the Effective Teacher Task Force's purpose was to determine how best to measure the quality of teachers and school leaders. The group publicly met 10 times and issued recommendations in August 2011. The task force recommended replacing the current pass/fail teacher evaluation system with five effectiveness levels determined, in part, by the results of VAMs. During the 2012 Legislative Session, House Bill 249, instituting a similar teacher evaluation system failed.

In April 2012, PED formulates and implemented a new teacher evaluation system making academic growth a key factor in teacher and principal evaluation. The new evaluation system was required as part of PED's waiver from No Child Left Behind. The NMTEACH workgroup, made up of teachers, administrators, union representatives, and other stakeholders, has since helped finalize a rule creating a new evaluation system. Fifty percent of the evaluation is based on a teacher's student achievement growth (35 percent on the SBA and 15 percent on other assessments); 25 percent on locally-adopted, PED-approved measures such as student surveys and short-cycle assessments; and 25 percent on observations of teaching. Fourteen school districts and 68 schools are piloting this system during the 2013 school year with statewide implementation scheduled to begin in SY14.

NEW MEXICO'S THREE-TIERED CAREER LADDER SYSTEM DOES NOT ALIGN PAY WITH STUDENT ACHIEVEMENT

Improving student achievement was a key policy goal of implementing the three-tiered system. The three-tiered system's founding legislation states student success for every child is the fundamental goal of New Mexico's education system. The three-tiered system was designed to help achieve this goal by attracting, retaining, and holding accountable quality and diverse teachers.

The three-tiered system is meant to be a progressive career system for teachers in which license level is a reflection of teacher ability, performance, and leadership. A level I license is a provisional license that allows beginning teachers to develop, whereas level II and level III represent teachers who meet and exceed department-adopted academic content and performance standards. Minimum salaries of $30 thousand, $40 thousand, and $50 thousand are tied to each licensure level to compensate teachers for their performance and abilities, as measured by the nine teacher competencies, as well as leadership roles taken at the school level.

Table 1. The Three-Tiered System's Licensure Levels

Licensure Level	Description	Minimum Salary
Level I	A provisional license that gives a beginning teacher the opportunity for additional preparation to be a quality teacher.	$30,000
Level II	A license for a fully qualified professional who is primarily responsible for ensuring that students meet and exceed department-adopted academic content and performance standards.	$40,000
Level III	A license for the highest level; for teachers that advance as instructional leaders in the teaching profession and undertake greater responsibilities such as curriculum development, peer intervention and mentoring.	$50,000

Source: NMSA 1978 22-10A-4

State law requires teacher evaluations to use a highly objective uniform statewide standard of evaluation (HOUSE). The Public Education Department developed nine key teaching competencies covering three areas of practice: instruction, student learning, and professional learning. Through the three-tiered system, teachers are evaluated against these competencies using local and state evaluation processes. A complete list of all competencies can be found in **Appendix B**.

State and local evaluations are essential pieces of the three-tiered system. Ascending levels within the three-tiered system is contingent upon evidence of satisfactory annual evaluations at the local level, as well as a satisfactory score on the Professional Development Dossier (PDD). While the local and state evaluations use the same competencies to evaluate teachers, only the local is conducted at the school site; state evaluations are conducted through the PDD, a collection of evidence of teacher performance that is reviewed externally.

Table 2. The Three-Tiered System's Evaluation System

Local Evaluation	State Evaluation (PDD)
Occurs regularly - every 1-3 years	Occurs when teacher applies for new license level
Teachers are evaluated on 3 strands of teacher competencies	Teachers are evaluated on 3 strands of teacher competencies
Evaluations are conducted by school administrators	Evaluations are conducted by independent reviewers
Focused on evaluating classroom practice	Focused on evaluating whether artifacts meet competencies
	Local Evaluations are taken into account

Source: NMSA and NMAC

Since 2009, nearly 6,000 teachers advanced to new licensure levels in the three-tiered system, receiving $59 million in mandatory salary increases. During that time 3,877 thousand teachers advanced from level I to level II, and 1,980 thousand advanced from level II to level III. Each advancement results in a minimum salary change of $10 thousand, or between a 25 percent to 33 percent increase in base pay.

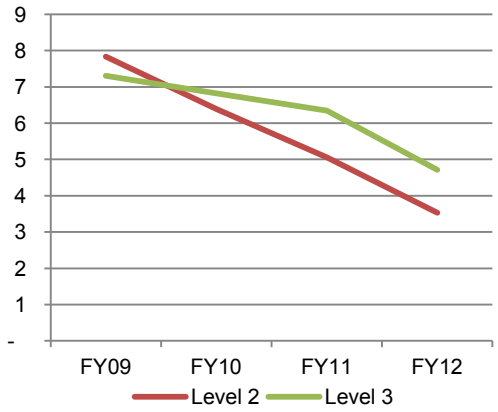
Table 3. SY09-SY12 Number of Teachers Ascending Licensure Levels

	SY09	SY10	SY11	SY12	Total
From level I to level II	904	1,278	786	909	3,877
From level II to level III	497	637	384	462	1,980
Total	1,401	1,915	1,170	1,371	5,587

Source: LFC Analysis of PED Data

Teachers advancing through the three-tiered system increasingly make up a larger proportion of classroom teachers than those grandfathered into their licensure level. As grandfathered teachers retire, the impact of the three-tiered system and PDD will become more apparent. The number of teachers grandfathered into the three-tiered system has declined steadily due to retirements and level II grandfathered teachers going through the PDD process to obtain a level III license.

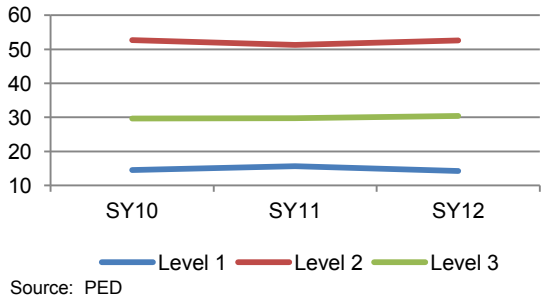
Chart 1. Grandfathered Teachers by Licensure Level
(in thousands)



Source: LFC Analysis of PED Data

Level III teachers comprise a larger and growing share of all teachers in the system. Between 2009 and 2012, the percent of level I teachers dropped from 15 percent to 14 percent of all teachers, while the percent of level III teachers increased nearly one percent, from 29.7 percent to 30.4 percent of all teachers.

Chart 2. Percent of Teachers by Licensure Level



Source: PED

In 2009, LFC staff recommended strengthening the three-tiered system to ensure teachers receiving large salary increases were producing better results for students. The LFC's 2009 evaluation of the three-tiered system and achievement gap found teacher licensure level did not significantly raise student achievement. Level III teachers generally had higher student achievement than teachers of other licensure levels, but they also were more likely to have a population of students more likely to succeed. The report recommended PED consider developing a bonus pay-for-performance pilot program. The report also called on PED to form a workgroup to evaluate proposals such as requiring more evidence of student performance in PDD submissions and teacher evaluations and establishing goals for expected gain on the SBA in grade levels and content areas.

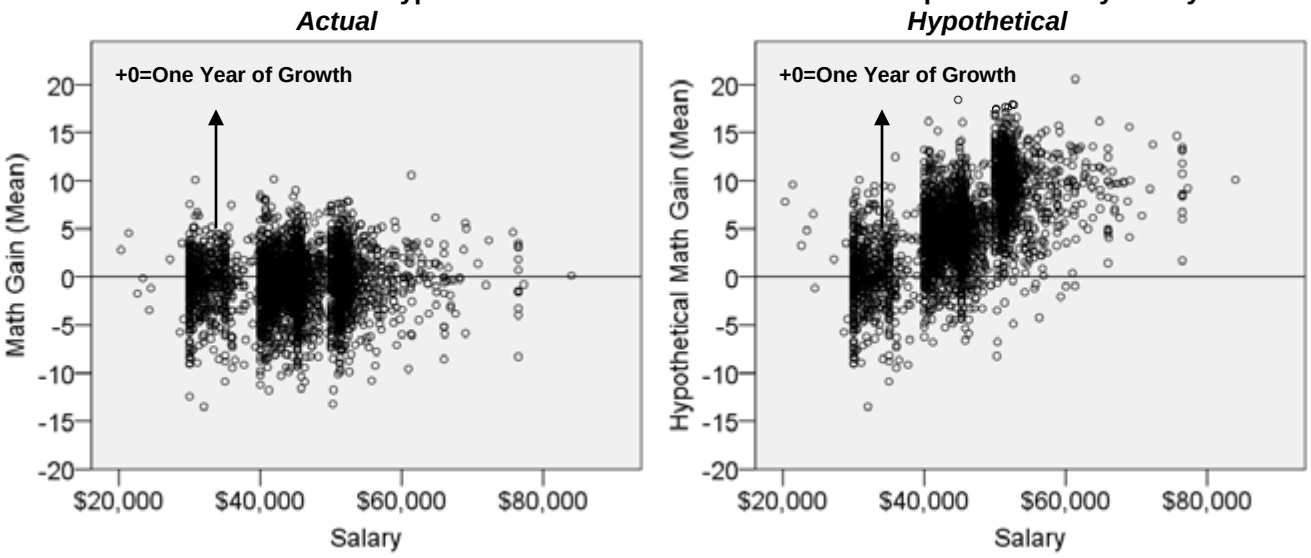
The state has not established expectations for student achievement across level I, II, and III teachers as part of evaluation systems. Competencies used in the state and local evaluations of the three-tiered system include examples of student performance, but the evaluations have no expectations for the performance of all students, particularly on standardized tests. When the three-tiered system was established, the SBA was new and lacked longitudinal information. As a result, student data was not incorporated into the evaluations. Teachers with different license levels achieve similar student performance, and a majority of New Mexico teachers do not feel the state evaluation process identifies effective teachers.

Student achievement is not a robust element of the current three-tiered system. To advance tiers through the PDD, teachers must submit and analyze student work. Teacher completing the PDD select examples of low, medium, and high-level student work and submit written reflections on that work. However, the 2007 joint report on the three-tiered system by the LFC, LESC and OEA stated, "These requirements focus primarily on describing or documenting student achievement, while involving no direct, explicit consequences – whether rewards or sanctions – for teacher based on the achievement of their students." Furthermore, the requirements in PDD only provide a picture of student performance for a few students, which are not easily compared across classrooms, schools, and districts. Student achievement data from statewide assessments like the SBA are comparable across the state.

As a result of no expectations for student performance, teachers across licensure levels produce similar student achievement results, despite large differences in pay and cost to taxpayers. Teachers in higher tiers generally produce better outcomes for students, but these differences are small and can often be accounted for by other factors. For example, after accounting for rates of English language learners (ELL), special education students (SPED), students participating in free and reduced-priced lunch (FRL), and ethnicity, the differences between tiers are further diminished.

The following scatter plots show how New Mexico teachers perform based on the average test score gain of their students in math as measured by the SBA and a hypothetical situation if more highly paid teacher were providing significant gains for students. Results for reading SBA scores are similar. A gain of zero represents a full year of academic growth for a student. For example, if a student scored a 40 on the 3rd grade reading SBA in 2011, and the following year scored a 40 on the 4th grade reading SBA, that student has maintained their proficiency level and met the challenge the 4th grade test provided. In chart 4, the scatter plot entitled *Actual* shows clear clusters of leveled teachers based on their minimum salaries. Each level achieves similar student achievement. If higher licensure levels were associated with higher student achievement, the scatter plot would look like the scatter plot entitled *hypothetical* in chart 4.

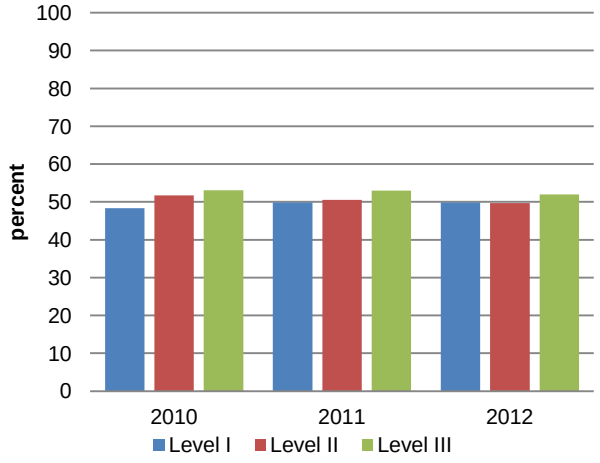
Chart 3. Actual and Hypothetical Math Student Academic Gain per Teacher by Salary



Source: LFC analysis of PED data

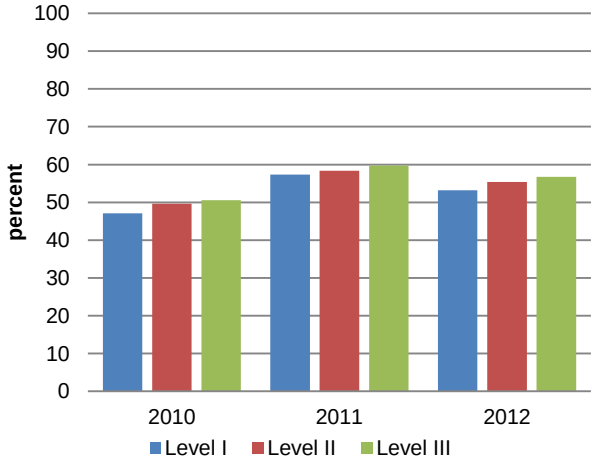
Based on student gains on the math and reading SBA, small differences exist between licensure levels. For example, in 2012, 52 percent of level III math teachers' students obtained a year's worth of growth, while 50 percent of level I and level II teachers' students accomplished the same feat. Even in years when the teachers with different licensure levels produced more disparate results, such as 2010, the difference between the percent of students who gained in math and reading was 4 to 5 percentage points, and significant differences did not exist among all tier levels. In 2010, 48 percent level I teachers' math students gained on the SBA and 53 percent of level III students gained on the assessment.

Chart 4. Percent of Math Students Obtaining One Year of Growth Per Tier (as measured by SBA gain scores)



Source: LFC Analysis of PED data

Chart 5. Percent of Reading Students Obtaining One Year of Growth Per Tier (as measured by SBA gain scores)



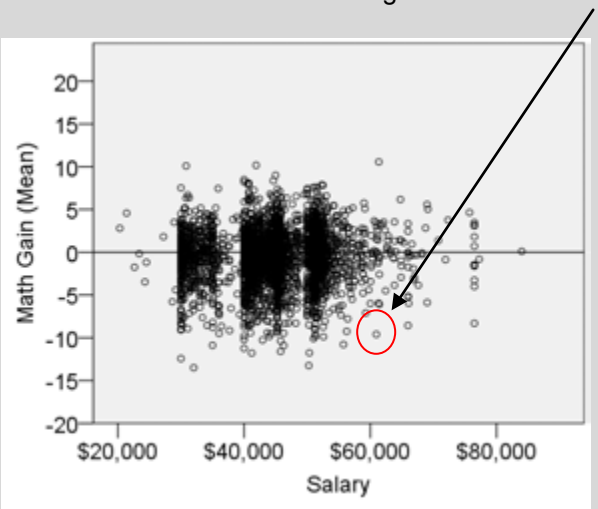
Source: LFC Analysis of PED data

Student demographics can explain part of the difference between licensure level performances. Level I teachers are more likely to teach students in poverty. Nationwide research and studies by the LFC found the obstacles facing students in poverty can affect achievement. For example, according to the LFC's program evaluation on developing early literacy in New Mexico, "Gaps persist in achievement between ethnicities, but the biggest differences are strongly associated with socioeconomic status and English language acquisition levels." Analysis from that study showed a 50 percentage point difference in reading proficiency levels between ELL and FRL students and their non-ELL, non-FRL peers. Teachers with a higher proportion of these students, such as level I teachers in New Mexico, may have lower test scores and fewer students showing gains on the SBA as a result.

Teachers in each licensure level perform similarly in moving their students up a proficiency level. Helping students increase a proficiency level on the SBA, such as an increase from a nearing proficiency rating one year to a proficient rating the following year, is uncommon: fewer than 20 percent of students moved up at least one proficiency level for math, and fewer than 25 percent moved up at least one proficiency level for reading during the last three years. Like the percentage of teachers' students who gain on the SBA, the difference between each licensure level of teachers is small, varied, and is not always statistically significant between tiers.

An example of a consistent poor return on investment

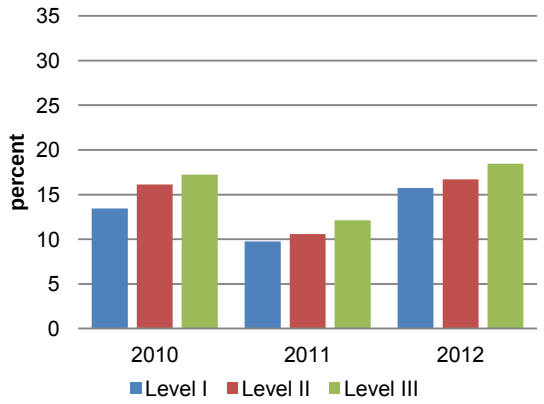
In 2012 an elementary teacher with over 30 years of experience was paid $60 thousand dollars. Their class of 15 students averaged a loss of 9.6 math SBA points compared to the previous year with over 70 percent of the class going down at least one proficiency level. The previous year the same teacher's class lost over 10 SBA points compared to the previous year. Losses for this teacher were also seen in reading SBA scores.



Source: LFC Analysis of PED data

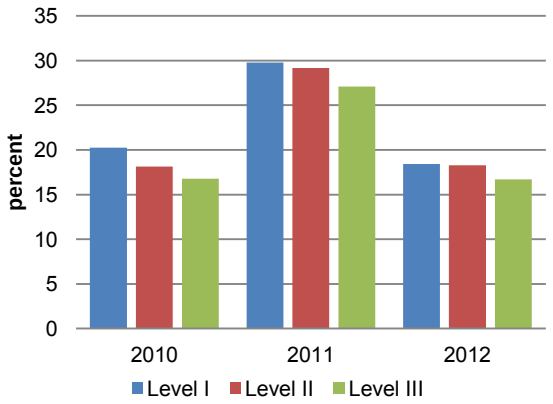
In math, students are losing more ground than they are gaining. Percentages for students moving down at least one proficiency level in math are higher than the percent of students moving up at least one proficiency level. Up to 30 percent of students lose at least one proficiency level in math, while less than 20 percent moved up at least one proficiency level in the last three years. Students are likely losing a proficiency level due to ineffective teaching and a lack of interventions.

Chart 6. Percent of Students Moving Up One Proficiency Level (Math)



Source: LFC Analysis of PED data

Chart 7. Percent of Students Moving Down One Proficiency Level (Math)

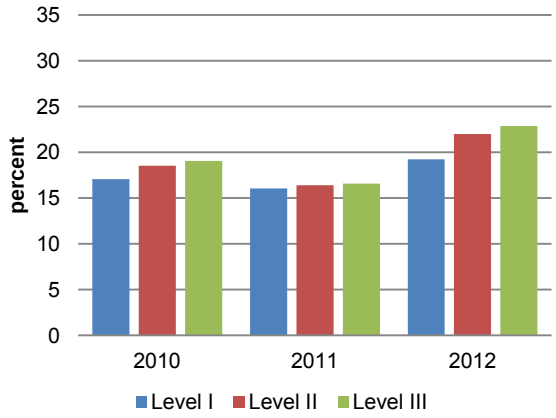


Source: LFC Analysis of PED data

In 2012, level III teachers were able to drive 2 percent more of their math students and 4 percent more of their reading students to higher proficiency levels. In 2011, level III teachers were able to move 1 percent more of their reading students up a proficiency level and 2 percent more of their math students up a proficiency level.

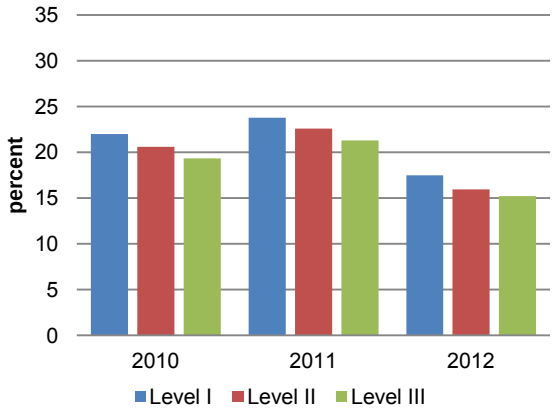
A smaller share of level III teachers' students move down a proficiency level, but the difference between them and level I teachers' students is small. In 2012, 17 percent of math teachers' students went down a proficiency level, while 18 percent of level I teachers' students moved down a level. Level II teachers had the same percentage of students decrease a proficiency level as level I teachers. Amongst all licensure levels and all years, more students move down a proficiency level than up. Almost 30 percent of a teacher's students moved down at least one proficiency level for math, and almost 25 percent of a teacher's students moved down at least one proficiency level for reading, during the last three years.

Chart 8. Percent of Students Moving Up One Proficiency Level (Reading)



Source: LFC Analysis of PED data

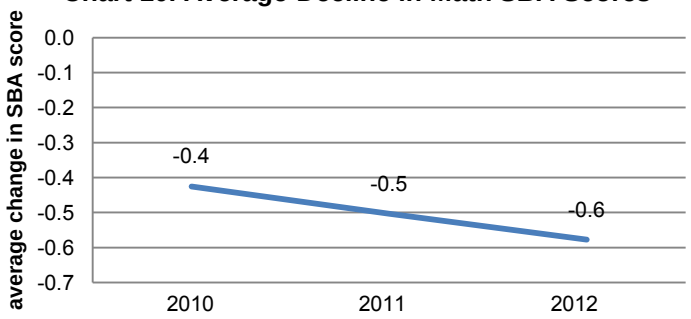
Chart 9. Percent of Students Moving Down One Proficiency Level (Reading)



Source: LFC Analysis of PED data

Math teachers' students across all teachers decline year-over-year. Math students in the state averaged a decline of 0.4 points on the math SBA in 2010 and a decline of 0.6 points on the assessment in 2012.

Chart 10. Average Decline in Math SBA Scores



Source: LFC Analysis of PED Data

Without more differences in student outcomes, the three-tiered system acts as an expensive proxy for paying teachers based on their education and experience. Licensure level, education level, and years of experience are strongly related amongst New Mexico teachers. This relationship is driven by PED's requirements for submitting a PDD, which include minimum years of experience and education level. A bachelor's degree is the highest degree obtained by 69 percent of level I, while 54 percent of level II teachers have a bachelor's as their highest degree completed. Of level III teachers, over half have earned a master's degree or higher. Similarly, level I teachers have fewer years of experience, 1.6 years, compared with level III teachers, 12.3 years.

The local evaluation system does not differentiate between high and low-performing teachers or focus on student achievement. Evaluation requirements need strengthening to align with common core, use student data, assess teacher effectiveness, and improve satisfaction among teachers. The competencies and requirements used in the state's PDD and local evaluation system have not been reviewed and changed since implementation and could be updated to better reflect current education research. Much research has been published regarding effective teaching since 2003, and a review and update of these competencies is needed to ensure the evaluation's effectiveness.

The PED, through their Effective Teaching Task Force and NMTEACH Task Force, has developed a new rule that changes the local evaluation structure for teachers. The new evaluations are broken down into three major components: 50 percent of the evaluation is based on a teacher's student achievement growth; 25 percent on locally-adopted, PED-approved measures; and 25 percent on observations of teaching. The new evaluations will be fully incorporated for all teachers during the 2014 school year.

Effective and explicit use of student data is only a small part of the current local evaluation structure. The performance evaluation system does not include a clear standard of practice for data use, particularly assessment data. Strand 2, competency 5 states "Uses information gained from ongoing assessment for remediation and instructional planning," but does not suggest how a teacher goes about using data effectively, and more precise expectations are needed to drive teacher development.

Not all teachers are annually observed and evaluated on their classroom practice. While level I teachers must have their classroom practice evaluated each year, level II and level III teachers only receive such evaluations every three years. Between these evaluations of their classroom practice, level II and III teachers use progressive documentation to provide formative information on performance. This progressive documentation, which does not require classroom observation, results in a summative performance evaluation every three years, which closely resembles a level I annual evaluation and includes classroom observations.

The current pass-fail evaluation system does not provide nuanced feedback or identify of teachers' ability. New Mexico uses a binary evaluation system that rates teachers as satisfactory or unsatisfactory. A recent study by the New Teacher Project, entitled *The Widget Effect*, analyzed 12 districts in four states to find 99 percent of teachers in districts using binary evaluation systems are rated effective. The study concludes this is problematic because excellence goes unrecognized and professional development opportunities and support cannot be properly targeted toward teachers who need it.

The New Mexico Effective Teaching Task Force made a similar finding in its final report; "Research indicates that multiple levels of effectiveness are needed in order to provide a mechanism for distinguishing average work performance from truly outstanding work performance."

Other measures not included in the current teacher evaluation system or PED's new teacher evaluation system are needed to assess complete teacher performance. While student achievement is integral to determining the effectiveness of a teacher, other measures must be incorporated into teacher evaluation systems to fully capture a teacher's performance. The Bill and Melinda Gates Foundation's Measures of Effective Teaching (MET) study suggests a fair and reliable teacher evaluation system includes the following five measures:

- Student achievement gains on assessments,
- Classroom observations and teacher reflections,
- Teachers' pedagogical content knowledge,
- Student perceptions of the classroom instructional environment, and
- Teachers' perceptions of working conditions and instructional support at their schools.

The New Mexico Effective Teaching Task Force came up with similar recommendations for multiple measures to be included in a new teacher evaluation system, noting "Effectiveness levels should only be assigned after careful consideration of multiple measures, including student achievement data, observations, and other proven measures selected by local districts from a list of options approved by New Mexico's Public Education Department." Because local districts select many of their multiple measures from a menu of options, some of measures might not be included in the evaluation.

Competencies need to be updated to align with the common core standards. PED calls the standards, which will be phased-in over three years and in full effect by 2015, "A different approach to learning, teaching and testing engenders a deeper understanding of critical concepts and the practical application of knowledge." Given the substantial change to the standards, testing, and expectations for teaching, the current competencies will need to be adjusted accordingly.

PED has recognized the shortcomings of the local evaluation system; however the proposed replacement requires a statutory change and is not linked to the three-tiered system. Through two task forces, the PED has designed and begun implementation plans for a new local evaluation system. The system includes the use of multiple measures and allows districts to determine which assessments they will use to evaluate teachers in addition to the SBA. This conflicts with HOUSE, which articulates the way teachers are evaluated must be uniform across the state. If each district to choose part of its own evaluation system, teacher evaluation will not be uniform across the state. A change in legislation is necessary to implement their rule.

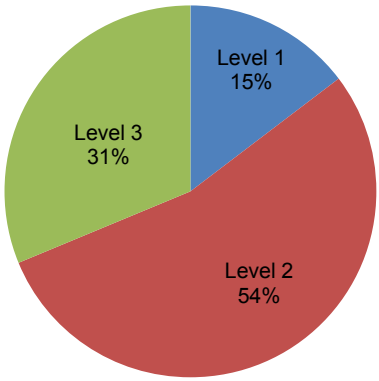
The professional development dossier does not effectively screen teachers for advancement, resulting in ineffective teachers receiving large pay increases. Teacher performance within licensure levels varies greatly because the state evaluation's PDD process does not effectively screen for high performing teachers. As a result, high and low performing teachers exist at each licensure level. The lack of clear and consistent performance among teachers in each licensure level shows the PDD process rewards teacher experience and education and not a teacher's impact on student achievement.

The PDD fails to effectively differentiate performance among teachers advancing licensure and pay levels; PED has not fixed this important part of the system. A 2009 study conducted by the LFC noted little difference between grandfathered teachers and teachers in licensure levels who passed the PDD. The report recommended using student achievement to drive the PDD process, but PED has yet to make student achievement a major factor in ascending licensure levels. The findings from the 2009 LFC report remain relevant to the current situation in the three-tiered system.

Over 90 percent of teachers who submit a PDD advance a level and receive large pay increases. Passage rates for teachers seeking level II and level III licenses indicate the system is providing raises for a vast majority of teachers who meet the experience and education criteria, rather than reaching an expected level of student performance. In 2012, 90 percent of the level II teachers who applied to move up to level III licensure were successful.

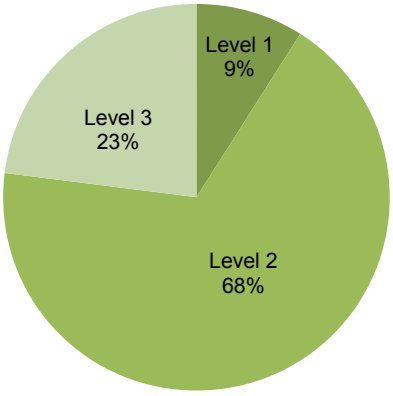
The differences in teacher performance within licensure levels vary greatly, suggesting the PDD process and competencies do not properly screen for teacher effectiveness. The PDD process does not do enough to focus and distinguish teachers based on student achievement, and as a result large amounts of high and low-performing teachers exist in each licensure level. In SY12, over two-thirds of high performing reading teachers had a level II license. Level II teachers made up the majority of underperforming reading teachers, but 30 percent of underperforming reading teachers were level III teachers. Underperforming teachers were identified as those performance is ranked in the lowest 16 percent of all teachers based on student performance on the SBA.

Chart 11. Proportion of Licensure Level Teachers as Percent of All Teachers



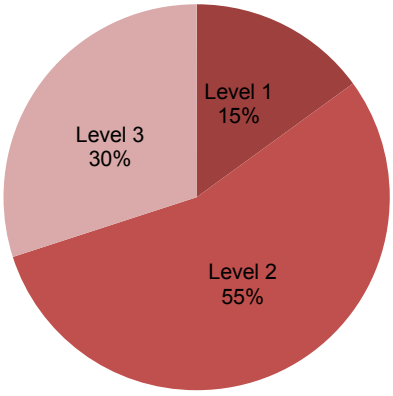
Source: PED

Chart 12. FY12 License Levels as a Proportion of High Performing* Reading Teachers



Source: LFC Analysis of PED Data
*High-performing teachers are those ranked in the top 16th percentile of all teachers

Chart 13. FY12 License Levels as a Proportion of Low Performing* Reading Teachers



Source: LFC Analysis of PED Data
*Low-performing teachers are those ranked in the bottom 16th percentile of all teachers

Only level II teachers were over-represented among high and low-performing teachers when compared with the proportion of level II teachers statewide. Fifty-four percent of teachers have a level II license statewide, but 68 percent of high-performing teachers and 55 percent of low-performing teachers have a level 2 license. Level I and Level III teachers are under-represented given their statewide percentages of 16 and 40 percent, respectively.

In SY12, the state awarded nearly $3 million in mandatory minimum salaries to low performing level-III teachers. Two-hundred eighty level III teachers ranked in the lower third of all valid math and reading teachers in moving students to a year or more of growth. In reading, this meant less than 40 percent of students achieved a year of growth; in math, ineffective level III teachers moved less than 30 percent of students a year's worth of growth. Seventy of the nearly 300 teachers were ranked in the lower third for both reading and math. Further, nearly 30 percent of the lowest performing teachers in the state had a level III license in SY12. Twenty-nine percent, or 271, of the bottom quartile of reading teachers have a level III license.

Table 4. Share of Teachers in Each Quartile of Teacher Performance, Based on SY12 SBA Reading Results

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Percent	Count	Percent	Count	Percent	Count	Percent	Count
Level 1	15%	142	13%	123	13%	119	11%	106
Level 2	55%	506	55%	534	54%	501	54%	507
Level 3	29%	271	33%	317	33%	306	35%	327

Source: LFC Analysis of PED Data

Two hundred sixty level III teachers are paid the maximum mandated salary but perform amongst the lowest quartile of math teachers. Twenty-six percent of level II teachers and 25 percent of level I teachers are in the lowest quartile of performers.

Table 5. Share of Teachers in Each Quartile for Teacher Performance, Based on SY12 SBA Math Results

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Percent	Count	Percent	Count	Percent	Count	Percent	Count
Level 1	15%	131	14%	124	16%	142	14%	125
Level 2	56%	495	56%	498	52%	462	52%	464
Level 3	29%	260	30%	266	32%	283	34%	298

Source: LFC Analysis of PED Data

Thirty-six percent of teachers disagree or strongly disagree the PDD successfully identifies highly effective teachers. In 2009, over half of teachers disagreed with the same statement. Since 2009, fewer teachers believe the PDD process positively impacts their ability to improve student performance or the materials required to submit a dossier adequately measure and reflect their skills as a teacher. The percentage of respondents who 'agree' or 'strongly agree' decreased by 9 percent and 4 percent, respectively. Many teachers felt the PDD process needed to include observations and should be include teacher observation.

> *"PDD is an OK tool, but not completely effective. Consistently random observation of classes is a quick way to see if teachers are at least managing and doing what is required. SBA is an effective tool for math, science and English, but not to grade the whole school. Each subject needs a diagnostic for each subject and grade area that reflects the student's knowledge of that grade-level subject area."*
>
> ***-Respondent from LFC Survey***

Despite a 92 percent rate of passage, only 25 percent of teachers believe the PDD is scored objectively and consistently. The respondents expressed concern with the honesty of the teachers submitting the dossier. After explaining they were encouraged to cheat on the dossier, one survey respondent noted, "The process relies on the integrity of the individual. Some are more honest than others. The artifacts used for evidence are good indicators of teacher effectiveness if they are authentic." Other respondents noted some teachers used "fake data" to successfully pass the PDD.

PED rule allows out of state teachers to bypass the PDD process for advancement to higher licensure levels. Whereas teachers within the state are required to go through the PDD process for advancement to a higher licensure level, a teacher who moves to New Mexico after teaching in another state or country can be placed into level II or level III licensure without passing the PDD. The presentation of a dossier is not required for these teachers and the basis of placement is total amount of years required for level placement.

The state allows low-performing teachers to keep their license level because the state does not have a rigorous license renewal process. Grandfathered teachers, those who obtained a level II or III licensure without going through the state's PDD, continue to renew their licenses without passing the PDD. Many of these teachers are low-performing teachers with the highest mandated salary in the state.

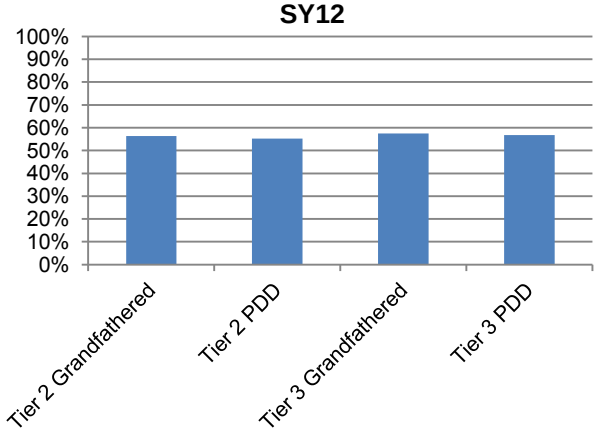
The requirement to renew licenses does not include demonstrated effectiveness. The requirement to renew licensure for New Mexico teachers is outlined in rule. Requirements include having demonstrated the competencies for the current level of licensure and meeting other requirements evidenced by annual evaluations. Therefore, a teacher can stay at level II for their career without completing the PDD. Although evaluations are taken into consideration for renewal, current evaluations do not use student performance measures based on SBA scores or other standardized student outcomes and therefore could allow underperformance over time.

A renewed level III license for a low-performing teacher

PED's inability to identify ineffective teachers and the lack of measures of effectiveness in license renewal has a detrimental effect on student achievement. For example, a level III teacher in a medium-sized school district in eastern New Mexico had over 20 per cent of their class lose a proficiency level in math for SY10 and the class averaged a two-point loss on the SBA. In SY11 this teacher's performance worsened significantly as over 60 p ercent of their class lost a proficiency level in math and on average lost eight points on the SBA. This teacher's level III license was renewed in 2012 despite this poor performance. Unfortunately, in SY12 this teacher's performance worsened again with over 70 percent of their class losing at least one proficiency level and the class, on average, losing 11 points on their SBA score compared with the previous year.

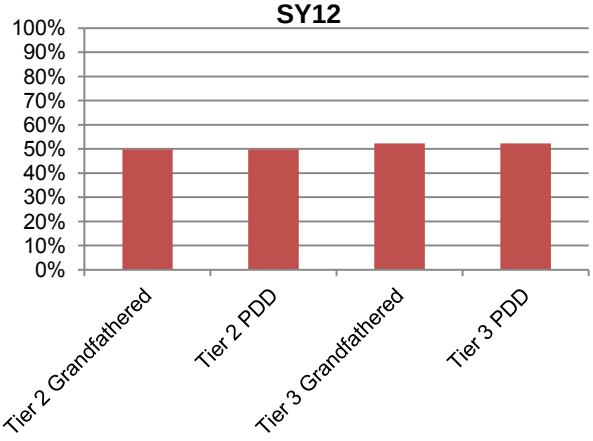
The state grandfathered thousands of teachers into higher salaries without a dossier, and does not require demonstrated effectiveness upon renewal of license. Students of PDD passers and those grandfathered into the system perform similarly, suggesting the three-tiered system has been unsuccessful in driving student achievement. The PDD is intended to provide sufficient evidence a teacher is qualified to advance to a higher licensure level and costs teachers hundreds of dollars to apply. However, in SY11 and SY12 there were no significant differences in student achievement between PDD and grandfathered teachers.

Chart 14. Students Obtaining One Year of Reading Growth Per Level In SY12



Source: LFC Analysis of PED Data

Chart 15. Students Obtaining One Year of Math Growth Per Level In SY12

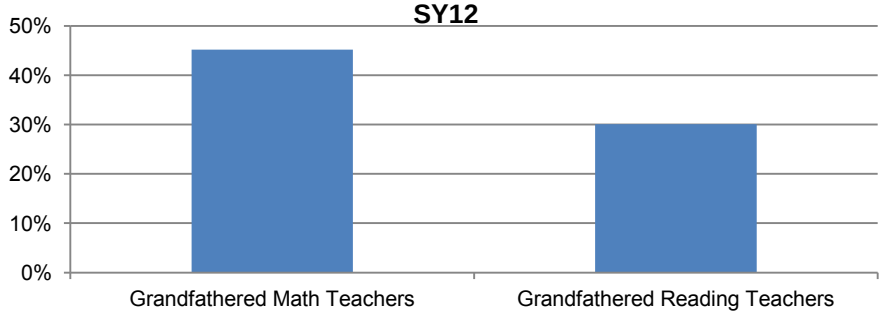


Source: LFC Analysis of PED Data

Grandfathered teachers show no significant differences in students moving up or down in proficiency or in raw gain score when compared to teachers who passed the PDD. Similarly, comparing the percentage of students with a year of growth on the reading and math SBA showed no significant difference between PDD passers and grandfathered teachers. On average level II and level III PDD teachers have 11 years of experience whereas grandfathered teachers at the same levels have 14 years of experience. After accounting for years of experience PDD and grandfathered non-PDD teachers perform similarly.

Nearly 50 percent of grandfathered teachers are in the lowest half of teachers when measured by student achievement. Forty-five percent of grandfathered reading teachers and 50 percent of grandfathered math teachers fall into the lowest half of teachers when measured by student performance. Many students in the classrooms of grandfathered teachers do not grow one year on the SBA in math or reading. Less than half of 126 grandfathered math teachers' students obtained a year's worth of growth, and 182 grandfathered reading teachers achieved the same feat.

Chart 16. Grandfathered Teachers with Less than Half of Students Obtaining a Year's Worth of Growth by Subject SY12



Source: LFC Analysis of PED Data

The three-tiered system offers a framework to align resource allocation to performance, but student achievement must be better incorporated into the process. The three-tiered system offers a framework for the state to align resources with results. The system offers significant salary increases and a competency structure; if modified, this system could drive student achievement across the state. A modified three-tiered system could serve as a way to strategically reward the state's best teachers and provide strategic interventions for struggling teachers.

The three-tiered system successfully retained teachers in New Mexico schools. A 2007 LFC, LESC, and OEA study found nearly 64 percent of teachers believed the three-tiered system helped with recruiting and retaining teachers. The study, which compared data from 2001 and 2007, found fewer teachers were leaving the profession within their first three years and fewer teachers overall were leaving to take positions in other states or outside of the teaching profession.

PED now captures student achievement data that could be incorporated into the PDD to make it more robust. PED now has many years worth of student achievement data linked to teachers to articulate their effectiveness. For example, SBA data can now connect students, demographic information, and teachers. Student achievement should be central to any changes to the three-tiered and teacher evaluation systems, but the state should move forward knowing that restrictions to the way data is currently collected and flaws in statistical models measuring student achievement will need to be addressed before the system is implemented.

Opportunity exists to incorporate demonstrated effectiveness into passage between tiers and renewal and better align funding and results. Incorporating demonstrated effectiveness in the form of student achievement could help provide incentives for better performance and deliver a higher return on investment for teachers as higher levels in the system. Both systems currently make student achievement a part of their systems but fail to make clear expectations about the amount of student achievement expected and the metrics to be used to measure student

achievement. Teacher evaluation systems should evaluate teachers on clear criteria such as student testing gain scores, whereas the three-tiered system could incorporate value-added models to identify the highest performing teachers in the state as one criterion for advancement. For example, forty-six percent of teachers with level II licenses who have been teaching for the last three years have averaged at least a year of growth for students in two of the last three years. A system that prioritizes these teachers for movement up to level III or renewal of their level II license would begin to align resources with results. This is just one example of how student achievement could be used to make student achievement a more robust element of the three-tiered system.

Student performance could then become the foundation for the way teachers and districts are compensated. By making licensure level more dependent on student performance and aligning the Training and Experience (T&E) index to the three-tiered system, the state can allocate resources to districts and teachers making the most impact in driving student achievement.

The system could allow exceptionally impactful teachers to ascend to level III licensure without submitting a PDD. Allowing high-performing teachers to move more freely in the three-tiered system will improve student achievement amongst teachers. Teachers who have proven themselves among the best educators in the state consistently and are not already a level III teacher could have a way to bypass the traditional PDD process and become a level III teacher based on their success in driving student achievement. A strengthened PDD is essential for other teachers in non-tested grades and subjects.

Recommendations:

The state should set performance expectations for licensure levels.

The Legislature should:

- Replace the PDD and establish the effective teaching portfolio (ETP) as part of the licensure advancement application with new requirements. The ETP will have new requirements strengthening expectations for student achievement. The legislature should adopt the following requirements for licensure advancement:
  Advancement to any licensure level: three years of classroom teaching at Level I before advancement; three years of satisfactory annual local evaluations; and superintendent approval of advancement and verification of submittal information.
  Advancement from license Level I to Level II: require one year of mentor program; meet performance expectations as demonstrated through an ETP – OR – qualify through Performance Ranking: Level I teachers ranked in the highest 50 percent of all level II teachers for three consecutive years can bypass ETP requirements and be promoted to level II.
  Advancement from license Level II to Level III: meet performance expectations as demonstrated through ETP – OR – qualify through Performance Ranking: Level II teachers ranked in the highest 50 percent of all level III teachers for three consecutive years can bypass ETP requirements and be promoted to level III;
- Modify statute to clarify that public school remedy for non-performance includes non-renewal of contract, or other action (suspension or termination) in accordance with other existing due process laws.

PED should:

- Create a Value-added model to estimate teacher effects on student performance.

The state should have a licensure renewal process that uses student achievement as a primary determining factor in the process.

The Legislature should:

- Create the following licensure terms for Level I, II and III licenses.
  Level I licenses should have a five-year license term and teachers must submit for advancement after three years; Level II and Level III licenses should have an eight-year term and teachers must submit for renewal after six years.

- Create the following requirements for license level renewals for Level II and III licensure holders:
  Level II requirements for renewal: satisfactory score on effectiveness evaluation for most recent three year period; satisfactory score on student achievement portion of ETP –OR – VAM ranking of 'meets expectations' within licensure level (Level II); a teacher failing to meet renewal requirements within license term may apply for a provisional Level II license and demonstrate satisfactory performance within two years.
  Level III requirements for renewal: satisfactory score on effectiveness evaluation for most recent three year period; satisfactory score on student achievement portion of ETP –OR- VAM ranking of meets expectations within licensure level (Level III) statewide; a teacher failing to meet renewal requirements within license term may apply for a provisional Level III license and demonstrate satisfactory performance within two years.

The state should update teacher competencies to incorporate student achievement, and reflect recent research and common core standards.
The Legislature should:

- Replace the PDD and establish the effective teacher portfolio (ETP) as part of the licensure advancement application. The ETP should include overhauled competencies that reflect current research, are updated to reflect the new common core standards, and provide evidence of effective teaching practice. New teacher competencies should focus on three areas:
  1) instruction, professional development and student learning: instruction includes evidence of instructional plans, assessment techniques, use of data to inform practice, adaptation of teaching for diverse learners, classroom management, and implementation of state content standards;
  2) professional development includes evidence of meeting professional development goals, collaborating with other educators, parent involvement, or research and publication;
  3) student learning includes evidence of improved student achievement on PED-approved assessments using at least three years of data. Evidence of student learning should constitute at least 50 percent of the overall ETP score, which should be heavily rooted in student year-over-year gain scores on the SBA math and reading assessments.

PED should:

- Establish updated basic competency and effectiveness indicators for teachers, as part of the ETP, including setting new student achievement expectations for Level II teachers and more rigorous student achievement expectations Level III teachers.

The state should strengthen local evaluations to better drive student performance.
The Legislature should:

- Require principals to receive training at least once every two years to improve evaluation skills;
- Strengthen statutory requirements for a highly objective uniform standard of evaluation (HOUSE) for teachers by requiring the following:
  - Professional Development Plan by 40th day establishing the current year's performance goals, including measurable objectives for student performance. The goals should be based on Basic Competency and Effectiveness Indicators, the previous year's annual evaluation, and previous year's students' performance;
  - Performance Evaluation: Annual evaluations should be based on whether the teacher met or exceeded expectations on Basic Competency and Effectiveness Indicators, made satisfactory progress on professional development goals, and received satisfactory ratings from students and parents. Classroom observations from principals;
  - Performance Improvement Plan: Establish a structure to provide assistance to teachers not meeting expectations.

- Local Schools: Create policies and procedures to implement this section and authorize PED to approve additional options and measures for a local system of data collection for the annual teacher performance evaluation, including the use of peer observations.

PED should:

- Establish an effectiveness evaluation for career teachers (level II and III):
    - After three years of classroom teaching require an effectiveness evaluation to be conducted no later than the 40th day the following school year and include three-year summaries of progress meeting Basic Competency and Effectiveness Indicators; improving student achievement component should count for no less than 50 percent an overall rating;
    - The summative effectiveness evaluation includes a cumulative assessment of a teachers' effectiveness at improving student achievement over time, as measured by PED expected student performance growth targets on the ETP. Performance expectations should be aligned with the three-tiered licensure levels, and subject and grade level standards; public schools may award teachers with successful effectiveness evaluations multi-year contracts not to exceed the equivalent term of a contract of the district's superintendent. Public schools may use the results of the effectiveness evaluation to make employment decisions, in accordance with other provisions of law.
- Require professional development plan by the 40th day establishing the current year's performance goals, including measurable objectives for student performance. The goals should be based on updated basic competency and effectiveness indicators, the previous year's annual evaluation, and a previous year's students' performance.
- Provide that a performance evaluation be conducted annually, and be based on whether a teacher exceeds, meets, or does not meet expectations on basic competency and effectiveness indicators, professional development goals, and satisfaction from parents. All teachers must be observed by principals three times a year.

WHEN USED APPROPRIATELY, VALUE-ADDED MODELS CAN IDENTIFY EFFECTIVE TEACHERS AND DRIVE STUDENT ACHIEVEMENT

Value-added models are increasingly used across the country to evaluate teacher performance. As of October 2012, the U.S. Department of Education granted 33 states, including New Mexico, a waiver from some of the requirements of No Child Left Behind for changing their teacher evaluation systems to incorporate student data. Many proposals included use of VAMs for 50 percent of a teachers' evaluation rating. VAMs have the potential to inform stakeholders about teacher performance, but the volatility in these models warrants caution moving forward because of potential misclassification of teachers.

Depending on the demographic factors used, value-added models produce varied results. Some VAMs attempt to control for demographic factors and may use multiple years of scores on a handful of different assessments, while others do not. Value-added experts debate the meaning of these different models.

In 2012, Pearson Education, Inc. published a study comparing five different VAM teacher evaluation approaches and concluded VAM results are not definitive, and will depend on the model used. **Appendix C** provides a comparison of these five models. The study used data from a large school district in Texas and included data on demographic variables such as gender, ethnicity, English proficiency, special education status, and FRL.

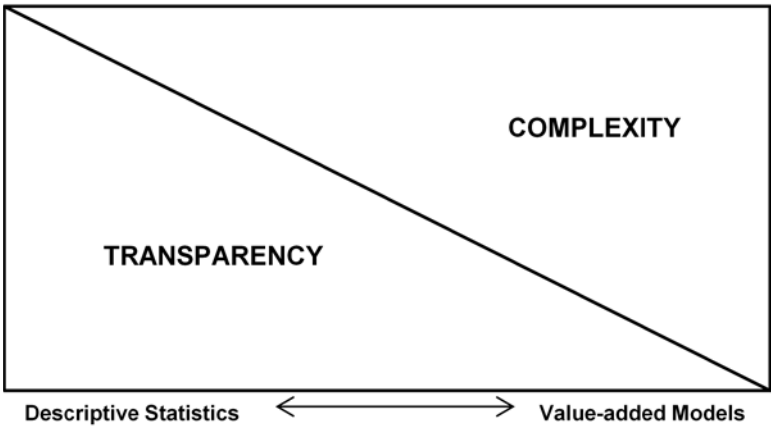
Chicago Teacher's Union Strike

In the fall of 2012, teachers in Chicago Public Schools, the nation's third largest school system, struck. Several reasons were cited for the strike, but, importantly, they included opposition to a system that would use student achievement data for 45 percent of teachers' evaluations. During the strike, researchers from 16 Chicago-area universities wrote an open letter to Chicago Mayor Rahm Emanuel warning against relying too heavily on VAMs in teacher evaluations. The New York Times later ran a rebuttal piece advocating for their use as prescribed. Chicago's students lost seven school days of school to the strike, which ended with a new contract agreement. The major tenants of this contract included a 17 percent pay raise for teachers and an evaluation system using VAMs for 30 percent, the minimum required by Illinois law.

Some VAMs produce unstable scores across years. All VAMs return different scores for the some teachers in different years. This could reflect a change in the effectiveness of a teacher between school years or the way a VAM score is calculated. Averaging VAM scores across a set number of years reduces this volatility, but such an approach does not help new teachers, teachers moving schools or grade levels, or teachers with small numbers of students from special populations.

More complex models are more consistent from year-to-year, but they can be difficult to interpret and expensive to implement. Multi-level VAM models, like the one PED uses for school grading, are complex and hard to explain to policymakers and stakeholders. The Pearson study indicates the importance of communicating the model and interpreting results to stakeholders and identifies this as an issue when choosing a model to use.

Figure 1. Understandability of Statistical Models



Additionally, the more complex the model, the more likely it will require custom programming; as a result, VAMs are increasingly being run by private for-profit corporations. For example, the software company SAS has gained the rights to a VAM approach developed by Dr. William Sanders in Tennessee and is now marketing the hosts data along with VAM analysis for districts and states.

Some value-added models adversely affect educators teaching certain populations of students. The Value-added Model Research Group at the University of New Mexico's College of Education used teacher and student data to determine scores for teachers from two different VAMs on teacher scores. One VAM incorporated only test scores (test-score only model), while the other compensated for contextual variables, such as poverty and English language fluency (student demographic model).

The two models placed teachers into one of five performance categories based on their students' performance.

Table 6. Performance Categories and Percentile Ranges of VAM Scores

Performance Level	Percentile Range
Ineffective	2nd percentile and below
Needs Improvement	2nd to 16th percentiles
Meets Expectations	16th to 84th percentiles
Highly Effective	84th to 98th percentiles
Exceptionally Effective	98th percentile and above

See **Appendix D** for a methodology behind the value-added models and the performance categories.

Mr. Wilson – Gifted Education Teacher

Mr. Wilson teaches at a large, urban district. In 2011, 97 percent of Mr. Wilson's students were classified as gifted. Twenty percent of his students qualified FRL and 42 percent were Hispanic. As a teacher of mostly gifted students, Mr. Wilson's VAM rating would vary depending on whether student demographic factors were included in the model.

How Different VAMs Affect Mr. Wilson's Status

	Test Score Only Model (no student demographic factors)		Student Demographic Model (includes all available student demographic factors)	
	Math	Reading	Math	Reading
1 year of data	Exceptionally Effective	Highly Effective	Ineffective	Ineffective
2 years of data	Highly Effective	Meets Expectations	Needs Improvement	Needs Improvement

Because Mr. Wilson specializes in teaching a high-performing group of students and improves their student achievement, models that do no t incorporate student demographics reflect his effectiveness. Mode ls that do incorporate student demographics penalize Mr. Wilson because his students are gifted and a relatively lower proportion of them are in poverty.

Teachers with high proportions of gifted students, students in special education, or poor students are affected by the model applied to them. Depending on whether the test-score only model or the student demographic model is used, some teachers with these populations of students have value-added scores showing them as high-performing or low-performing. For example, in a model without student demographics, teachers with high levels of FRL students have lower scores than a model including student demographics.

Table 7. Special Populations Adversely Affected When the Test Score Only and Student Demographic Models are Applied

Test Score only-model	Student Demographic Model
English-language learners	Gifted
Free Lunch	
Special Education C and D	

Source: UNM

Many teachers' ratings change depending on the value-added models used to estimate their impact on student achievement. After applying and comparing the results of both models, 77 percent of the teachers evaluated, stayed in the same performance category. The choice of VAM used affects the rating of 23 percent of teachers who moved from at least one performance category to another. Twelve percent increased at least one performance category and 11 percent decreased at least one performance category.

Mrs. Martinez – Special Education Specialist

Mrs. Martinez teaches 4th grade in a western New Mexico school district. In 2012, 68 percent of Mrs. Martinez's class was FRL. Mrs. Martinez has experience as a special education teacher and has shown the ability to drive their improvement, so nearly all of the students assigned to her class have IEPs. Half of her students are Caucasian. Mrs. Martinez's value-added scores are inconsistent across models.

How Different VAMs Affect Mrs. Martinez's Status

	Test Score Only Model (no student demographic factors)		Student Demographic Model (includes all available student demographic factors)	
	Math	Reading	Math	Reading
1 year of data	Ineffective	Needs Improvement	Meets Expectations	Highly Effective
2 years of data	Meets Expectations	Meets Expectations	Meets Expectations	Highly Effective

Teachers like Mrs. Martinez, with a high proportion of their students in special education, can be significantly affected by the type of model selected to evaluate them. Because of the population of students she teaches (nearly all special education), Mrs. Martinez could be judged as one of the state's very best teachers or one of the worst.

Value-added models are limited in what they can tell educators, the public, and other stakeholders. VAMs compare teacher performance to one another, making them a relative concept. One set of scores from VAMs do not indicate whether quality instruction is occurring in classrooms, only whether one teacher has students with higher achievement gains than another. Over time, VAM scores across multiple years can create a picture of absolute teacher performance that can be used for licensure advancement.

Value-added models used in other states can distinguish high-performing teachers from low-performing ones with some certainty, but cannot reliably distinguish between the middle-performing teachers. In New York, 77 percent of teachers earned a rating of "effective," meaning their students grew "equal to the average for similar students." The state's evaluation system gives teachers one of four overall ratings: highly effective, effective, developing and ineffective. Seven percent of teachers earned a rating of "highly effective," and 6 percent earned a rating of "ineffective."

Ms. Campos – Teacher of At-risk Students

Ms. Campos teaches 3rd grade in small, rural school district. Over 90 percent of her students are FRL and are classified as English language learners. Additionally, all of her students participate in special education and 100 percent of her students are Native American. As a teacher of this highly at-risk group of students, Ms. Campos' value-added rating in a VAM depends heavily on which model is applied.

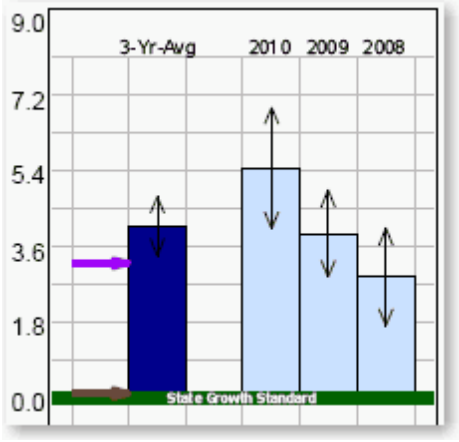
How Different VAMs Affect Ms. Campos' Status

	Test Score Only Model (no student demographic factors)		Student Demographic Model (includes all available student demographic factors)	
	Math	**Reading**	**Math**	**Reading**
1 year of data	Needs Improvement	Needs Improvement	Highly Effective	Highly Effective
2 years of data	Needs Improvement	Meets Expectations	Needs Improvement	Meets Expectations

Ms. Campos moves between one of the lowest performance categories to the highest. Using a model with no student demographic factors could discourage effective teachers from accepting positions in low-income schools.

A large proportion of scores from value-added models have a high degree of uncertainty. VAMs are unable to pinpoint the exact ranking of a teacher and instead provide wide estimates of performance. For example, in the sample below taken from the Tennessee Value-Added Assessment System, the teacher received a VAM score of 5.4 in 2010. However, the teacher's real score could be anywhere between approximately four and seven after taking statistical significance into account.

Chart 17. Example of Value-added Model Scores from Tennessee



Source: Tennessee Department of Education

Using the SBA, currently the only assessment appropriate for statewide value-added models, approximately 20 percent of teachers can be evaluated. Standardized tests are necessary inputs for VAMs. Because the SBA assesses only reading, math, and science, most students and teachers in the state cannot be evaluated under the model. This includes teachers of social studies, vocational education programs, art, and more. Even science teachers could be excluded from a model because the science SBA is only administered in grades 4, 7, and 11, and consecutive years may be needed to compute VAM scores. A district using additional standardized assessments could use a VAM on teachers within the district. Some VAMs use a school's aggregate score to evaluate these teachers.

Figure 2. Teachers in New Mexico

Figure 3. Teachers in New Mexico for whom we can compute VAM scores

Principal evaluations and some VAMs are identifying the same high-performing teachers. One study of principals in New York City found teachers rated more effective by their principals were more likely to also have high value-added test scores.

Some VAMs may not include many students due to mobility, test exemption, and absenteeism. VAMs use data from multiple years. If a student moves from one school to another during the school year or between school years or is absent for much of the school year, that student will not meet requirements to be considered as attending a full academic year. Furthermore, students who do not take the SBA or receive exemptions will not count in a VAM. Alternative methods exist for filling in missing scores in statistical modeling.

The use of VAMs can be responsibly integrated into the three-tiered system to identify teachers for advancement and bonus pay. VAM methodologies being developed for teacher evaluation and school grading could be leveraged to measure performance at tier levels and could be used as benchmarks in advancement between levels. Specifically, VAMs can be used to reward good teachers and identify poor teachers for professional development as part of a larger system of teacher evaluation.

VAMs can do a good job of identifying highly effective and highly ineffective teachers for rewards and interventions. Once teachers of certain populations that experience variability are identified and controlled for, research has shown VAMs do a good job of identifying very good and very poor teachers, but do not differentiate between teacher scores in the middle. Districts and states can reliably use this data to reward very good teachers and put them in roles allowing them to share expertise, such as a mentor. Similarly, identifying low-achieving teachers allows districts to strategically align resources.

School leaders have shown they use value-added data to drive professional development and improvement in classroom instruction. Research at Columbia University showed principals given teachers' VAM scores and performance data used the data to make personnel decisions and plan their teacher professional development. There were no formalized expectations for use of the data but principals, after being provided training on how to interpret scores, used the data in ways they found most advantageous for the school.

VAMs should not be used in annual local evaluations because of their limitations and complexity. Because VAMs do a good job of identifying very good and very poor teachers but do not do a good job of differentiating between teacher scores in the middle, the use of VAM in annual local evaluations could provide inaccurate reflections of teacher performance. Additionally, the use of multiple years of VAM data to boost reliability would hide recent changes in teacher performance. For example, if a teacher made a large gain between the previous and current year, this progress could be diminished by averaging the most recent score with multiple previous years of data. Additionally, a VAM score is a relative measure which provides information on how a teacher performs compared with other teachers. Teachers working toward definitive goals such as SBA proficiency scores of 40, an absolute measure, would not be provided useful information by a relative VAM score.

VAMs should be used cautiously. VAMs should not be applied to all teachers of students with test scores, because of special circumstances. For example, some special education teachers specializing in work with severely high-needs special education populations might not be appropriate to score on a VAM because of their students' limited ability to take the assessment. Similarly, teachers of exceptionally high-scoring students are less able to raise students' scores, and should not be subjected to VAM scores or their associated consequences. Exceptions for the small minority of teachers in these circumstances should be made when considering how best to implement VAMs.

Recommendations:

The state should take advantage of value-added models to

The Legislature should:

- Require PED to annually rank the performance of licensed teachers providing instruction in tested grades and subjects through a valid value-added model. Results will be provided only to public schools and the individual teachers. The results should benchmark performance relative to teachers statewide, within the district, the school, and license level by grade taught and overall. The results should be part of a teacher's personnel file, confidential, and only available for review by authorized personnel.
- Create a fast-track licensing process for teachers that receive the very highest statewide value-added scores; teachers consistently (3 years or more) receiving the highest scores could apply for level III licensure regardless of their current licensure status.
- Allow findings from future research by the Value-added Model Research Group at UNM, which aims to validate the accuracy of the VAM rankings and conduct research on special education population's effects on the state, to guide future decisions regarding the use of VAMs.

PED should:

- Use two different VAMs to obtain a composite score to help eliminate VAM biases for teachers of certain populations;
- Allow findings from future research by the Value-added Model Research Group at UNM, which aims to validate the accuracy of the VAM rankings and conduct research on special education population's effects on the state, to guide future decisions regarding the use of VAMs.

RESOURCE ALLOCATION AMONGST DISTRICTS AND SCHOOLS CREATES FUNDING DISPARITIES BETWEEN LOW-INCOME STUDENTS AND THEIR MORE AFFLUENT PEERS, WITHOUT DRIVING STUDENT ACHIEVEMENT

New Mexico directs nearly $200 million for higher teacher compensation through the T&E index in the public school funding formula. T&E index values are based on teachers' highest degree earned and years of experience. That index value is multiplied by student demographic and program units already generated in the formula. Funds generated by the T&E index in the funding formula compose up to 20 percent of a district's formula funding.

Table 8. Percentage of Formula Funding from the T&E Index

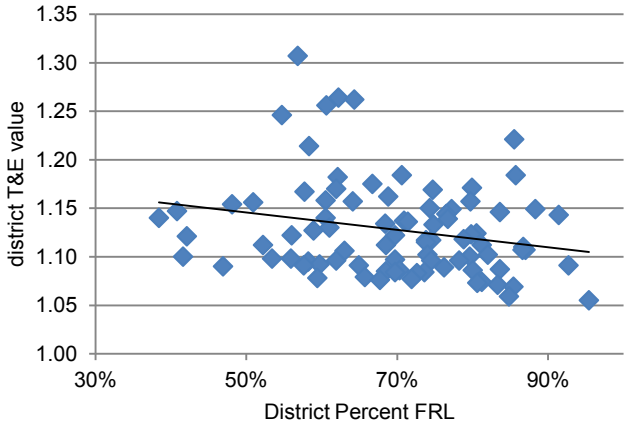
District	Share of Total Formula Funding
Carlsbad	19.8%
Tatum	16.6%
Texico	15.2%
Silver City	14.2%
Animas	13.7%
Ruidoso	12.9%
Cobre	12.8%
Artesia	12.7%
Tularosa	12.7%

Source: PED

The T&E index directs more funding to more affluent school districts and produces a questionable return on investment after factoring in poverty. The T&E does not recognize better performance by teachers and higher pay, but instead rewards relatively affluent districts for keeping teachers and sometimes requiring them to meet higher education requirements. Aligning the T&E index to a modified three-tiered system that focuses on student performance will allow the state to send resources to high-performing teachers and schools.

High-poverty, rural districts with the greatest needs generally have difficulty hiring experienced teachers with advanced degrees who increase the T&E index. As FRL participation levels increase, T&E values drop. Because poverty is also highly related to lower SBA scores, districts with lower T&E values sometimes tend to have lower SBA scores than their peers.

Chart 18. District T&E and Free and Reduced-Priced Lunch Levels



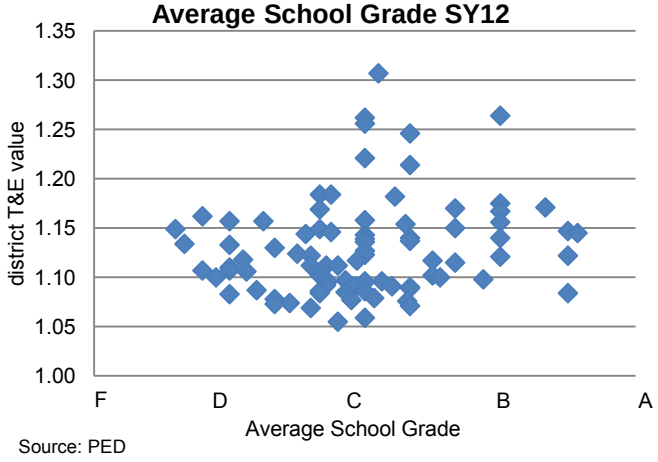
Source: PED

An independent study conducted for the Maddox Foundation also found "districts with a high T&E index tend to have relatively fewer at-risk students." The report recommended the state drop the T&E index from the formula and develop a categorical aid program focused on providing funds for high-need districts to improve teacher qualifications and recruit teachers in hard-to-staff schools and subjects.

The T&E index encourages higher education levels and more years of service, despite inconclusive evidence these factors increase student achievement. For example, some school districts, including those with some of the highest T&E values, mandate teachers earn a master's degree before a certain number of years serving with the district. No clear body of research links higher education level or more years of service with better student outcomes or achievement. A 2009 report from the Urban Institute states characteristics such as graduate education and experience are at best weak predictors of a teacher's contribution to student achievement. However many states, including New Mexico, continue to use education and experience as the basis for teacher compensation.

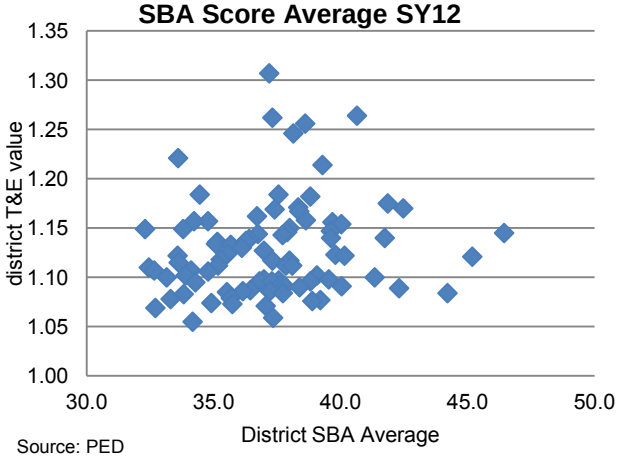
After accounting for district free and reduced-price lunch levels, T&E has no effect on district school grades, average test scores, or percent of students proficient. T&E has no relationship with district achievement as measured in numerous ways including using PED's school grading system and district SBA score average.

Chart 19. District T&E and Average School Grade SY12



Source: PED

Chart 20. District T&E and SBA Score Average SY12



Source: PED

Even districts with the very highest T&E values generally have average school grades. Of the five districts that earn the highest T&E values in the public school funding formula, half of them average C grades or less, and only one district averaged a B grade. Many districts that claim lower-than average T&E values have better school grades than many of the highest T&E earners; for example, Rio Rancho and Moriarty school districts, which each claim a 1.1 on the T&E index, had average school grades of 2.6 (C) and 2.9 (C), respectively.

Table 9. Districts with the Highest and Lowest T&E Index Values with Average School Grades and SBA SY12

District	Average School Grade	T&E Index	SBA Average
Tatum	2.1 (C)	1.31	38.4
Animas	3.0 (B)	1.26	40.9
San Jon	2.0 (C)	1.26	38.7
Carlsbad	2.0 (C)	1.26	39
Texico	2.3 (C)	1.25	38.6
Tucumcari	2.3 (C)	1.07	38.1
Jemez Mountain	1.6 (D)	1.07	34.1
Santa Rosa	2.0 (C)	1.07	38.8
Hatch	1.8 (D)	1.06	34.2
Lake Arthur	2.0 (C)	1	38.5

Source: PED

The T&E index is not aligned to the three-tiered system. The three-tiered system provides large salary changes not accounted for in the T&E index. For several years the LFC has noted, in its present form, the T&E index is not aligned to the three-tiered system. An evaluation of the public school funding formula conducted jointly by the LFC and the LESC, the New Mexico Effective Teaching Task Force final report, and the AIR funding formula study recommended better alignment of the T&E index with the three-tiered system. For example, in 2011, the LFC and LESC recommended the T&E index be replaced with an effective teacher index that assigns values to teachers based on licensure level, not education level and experience.

Forty percent of teachers in New Mexico charter schools are level I teachers, more than double the percent in Albuquerque Public Schools (APS), the district with the most charters. While the proportion of level I teachers varies throughout the state's regions, charters have between two to three times the proportion of level I teachers. Subsequently, their level II and level III teachers are lower than all regions throughout the state.

Table 10. Ratio of Teacher Licensure Level by Region SY12

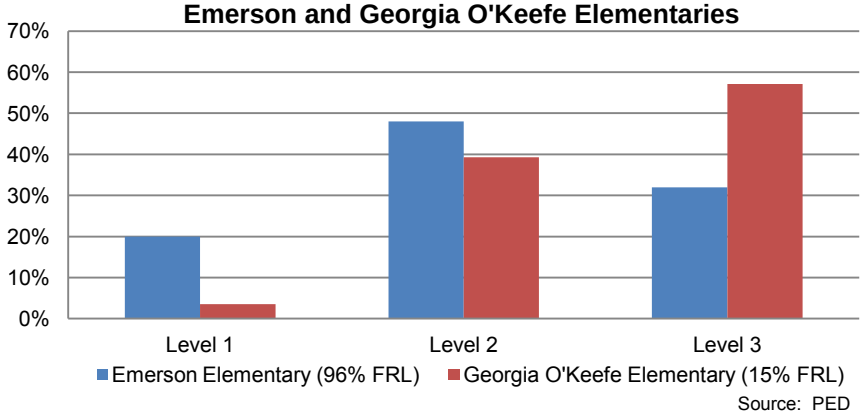
Region	Level I	Level II	Level III
APS	19%	44%	37%
NW	18%	47%	35%
NE	12%	52%	36%
SE	17%	44%	38%
SW	17%	44%	39%
Charter	40%	33%	28%

Source: PED

Level III teachers are more likely to teach in more affluent districts and schools. While districts do not have explicit policies to move more experienced and educated teachers to more affluent schools, students in poverty are more likely to have a less experienced, poorer-performing teacher. Recruiting and retaining high quality teachers in low-income schools is integral to ensuring students in poverty achieve academically.

Schools with higher poverty rates have teachers with lower licensure levels than more affluent schools. Based on an analysis of four school districts, Albuquerque, Las Cruces, Rio Rancho and Santa Fe, schools with higher proportions of FRL students have staffs with lower licensure levels. For example, in APS, Emerson Elementary and Georgia O'Keefe Elementary have similar staff sizes but different student populations: 96 percent of Emerson's students are FRL compared with 15 percent at O'Keefe. Over half of Georgia O'Keefe's teachers have a level III license and the school has only one level I teacher. Emerson's teaching staff is mostly level I and level II teachers.

Chart 21. Comparison of Staff Licensure Levels at Emerson and Georgia O'Keefe Elementaries



Source: PED

Teachers at high-poverty schools are paid less than those at low-poverty schools. Because high-poverty schools have more teachers with lower licensure levels, the average pay is lower. These districts do not have explicit policies placing level III teachers in more affluent schools.

Table 11. SY12 Average Salaries at High and Low-Poverty Elementary Schools in Selected Districts

District	Elementary School	Average Salary	Free and Reduced-Price Lunch Rate
APS	Emerson	$42,900	96%
	Georgia O'Keefe	$46,923	15%
Santa Fe	R.M. Sweeney	$41,503	87%
	Wood-Gormley	$49,506	19%
LCPS	Booker T. Washington	$43,714	90%
	White Sands	$46,385	30%

Source: PED

Research shows teachers at high-poverty schools are less effective than their counterparts in more affluent schools. A Duke University study of high and low-poverty schools in North Carolina found students in the high poverty schools are served by teachers with lower qualifications than those in lower poverty schools. Researchers found these qualifications were connected to higher student achievement. The study considered competitiveness of the teachers' undergraduate institution, teacher scores on licensure exams, national board certification, and years of experience.

Offering a mix of incentives to recruit and retain good teachers in high-poverty schools can work. Research finds creating incentives to get highly qualified and effective teachers to teach in high-poverty schools can work, but keeping effective teachers is more challenging. Financial incentives can recruit high-quality teachers and slightly decrease turnover in the short-term, but money does not work in the long-term to keep teachers at low-income schools: "Even when bonuses succeeded in drawing teachers to the poorest schools, such incentives could not compensate for the lack of support they encountered in these schools, which in turn contributed to the departure of many of these teachers."

Financial incentives attack part of the problem, but do not solve working conditions. Districts must find ways to incentivize the best administrators to lead high-poverty schools and give them added support.

Good results in challenging environments and poor results in advantageous settings

In SY12, Teacher X made $50 thousand a year and had a 100 percent of their students qualify for free and reduced-price lunch. Teacher X is an effective teacher moving more than half of the class up in proficiency level with no children losing a proficiency level. Students averaged a five point gain in SBA scores across students.

The same year, Teacher Z made $53 thousand a year in another district and has a 0 percent participation in the free and reduced-price lunch program. Teacher Z is a less effective teacher, having twice as many students in their class losing proficiency than gaining and with students averaging a two point loss in SBA scores across students.

Recommendations:

The state should align resources to allocate funding to districts in a way consistent with teacher pay scales that incorporate performance.
The Legislature should:

- Change the T&E index to an effective teacher index that rewards districts based on the number of teachers they have in each license level. This should be accomplished over a two-year time period: in year one, the current T&E index should be multiplied by only membership units in the formula, and in year two, the effective teacher index should be fully implemented.

The state should incentivize high-performance among teachers and provide incentives for teaching in high-need schools.
The Legislature should:

- Require only teachers meeting or exceeding expectations on annual performance evaluations receive state or district funded salary increases the subsequent year; and
- Consider a mechanism, possibly through the funding formula, to provide additional compensation to effective teachers (as measured by the new aforementioned teacher evaluation and three-tiered licensure system) to teach in Title I schools.

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

HANNA SKANDERA
SECRETARY-DESIGNATE OF EDUCATION

SUSANA MARTINEZ
Governor

November 13, 2012

Mr. David Abbey, Director
Legislative Finance Committee
325 Don Gaspar, Suite 101
Santa Fe, NM 87501

RE: Teacher Effectiveness Evaluation

Dear Director Abbey:

Thank you for the opportunity to respond to the draft evaluation on Effective Use of Test Data to Assess and Improve Teacher Evaluation in New Mexico. Please accept my compliments to your staff for their professionalism and collaborative approach throughout the evaluation process. As always, the Public Education Department (PED) is committed to continuous quality improvement, best practices, and positively impacting outcomes for all of our students in New Mexico.

The evaluation appears to be thorough and objective and points to a number of issues that we believe are necessary to implement a robust teacher evaluation system that will lead to a more effective workforce focused on ensuring students receive the education they need to thrive and survive in the 21st century. We are encouraged that the evaluation aligns with so many of the components of the newly developed evaluation system and in those minor areas where PED had concerns. We are pleased that the LFC has taken comment and made minor changes accordingly.

A key component of the LFC evaluation is the need to reward highly effective teachers, those that through thick and thin, make a difference in each student's life. I encourage the LFC to work toward changes in the funding formula that lead to these rewards, while holding these teachers to the highest standards for the benefit of all.

The exit conference between LFC and PED was held Wednesday November 7, 2012 and the draft report was discussed. The department does not have any recommended changes at this time. However, we note that several of the recommendations made in the LFC report are already being addressed as part of the teacher evaluation pilot currently operating in 68 schools across 21

districts. We look forward to working with your staff as the new evaluation system unfolds to ensure an evaluation process that is robust, fair and truly focuses on improving the teaching skills of all teachers.

Thank you again for the opportunity to comment on the evaluation.

Warm regards,

Hanna Skandera
Secretary-Designate
Public Education Department

HS/at

APPENDIX A: Project Information

Evaluation Objectives.

- Follow-up on LFC's 2009 three-tiered system report.
- Investigate T&E index and return on investment.
- Evaluate existing value-added model (VAM) methods and outputs.

Evaluation Procedures.

- Interviewed district and state-level administrators.
- Reviewed state, district, and school-level student performance data and student demographic data.
- Conducted an online survey of teachers that have gone through the three-tiered system.
- Reviewed and determined the impact of the current three-tiered system. Additional descriptive and inferential statistics, along with specific methodologies will be made available in a separate publication.
- Reviewed applicable laws and regulations; previous research reports; LFC file documents, including all available project documents; relevant performance reviews from other states; and performance measures.
- Evaluated VAM methodologies in collaboration with the Value-added Research Group at the University of New Mexico.

Evaluation Team.

Matthew Pahl, Lead Program Evaluator
Dr. Jon Courtney, Program Evaluator
Elaine Romero, Program Evaluator

The Value-added Research Group at the University of New Mexico calculated value-added scores for teachers using teacher and student data. The group conducted analyses using two value added models, which are described in detail in **Appendix C.** The Value-added Research Group included the following members:

Dr. Richard Howell, Dean of the College of Education, University of New Mexico
Dr. Richard Bowman, Chief Accountability and Strategy Officer, Santa Fe Public Schools
Dr. Kristin Umland, Associate Professor, Department of Mathematics and Statistics, University of New Mexico
Dr. James Selig, Assistant Professor, College of Education, University of New Mexico
Dr. Laura Kapitula, Assistant Professor, Department of Statistics, Grand Valley State University

Authority for Evaluation. LFC is authorized under the provisions of 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. LFC is also authorized to make recommendations for change to the Legislature. In furtherance of its statutory responsibility, LFC may conduct inquiries into specific transactions affecting the operating policies and cost of governmental units and their compliance with state laws.

Exit Conferences. The contents of this report were discussed with Secretary-designate Skandera and Senior PED staff on November 7, 2012.

Report Distribution. This report is intended for the information of the Office of the Governor; the 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.

Charles Sallee
Deputy Director for Program Evaluation

New Mexico Teacher Competencies for Licensure Levels I, II, and III
Assessment Criteria

New Mexico is one of the most diverse states in the nation, and this diversity is reflected in the strengths and needs of New Mexico's students. The ability of a highly qualified teacher to address the learning needs of all New Mexico's students, including those students who learn differently as a result of disability, culture, language, or socioeconomic status, forms the framework for the New Mexico Teacher Competencies for Licensure Levels I, II, and III-A Assessment Criteria Benchmarks.

1. The teacher accurately demonstrates knowledge of the content area and approved curriculum

Provisional Teacher - LEVEL I	Professional Teacher - LEVEL II	Master Teacher - LEVEL III
A. Utilizes and enhances approved curriculum.	A. Enhances and extends approved curriculum.	A. Contributes to the refinement and development of the approved curriculum.
B. Gives clear explanations relating to lesson content and procedures.	B. Gives clear explanations relating to lesson content and procedures.	B. Provides clear explanations relating to lesson content and procedures in multiple ways and is aware of knowledge and preconceptions that students can bring to the subject.
C. Communicates accurately in the content area.	C. Communicates accurately in the content area.	C. Communicates accurately in the content area and can create multiple paths to the subject matter.
D. Shows interrelatedness of one content area to another.	D. Integrates other subjects into the content curriculum.	D. Can articulate to students the interrelatedness of the disciplines.

2. The teacher appropriately utilizes a variety of teaching methods and resources for each area taught.

Provisional Teacher - LEVEL I	Professional Teacher - LEVEL II	Master Teacher - LEVEL III
A. Provides opportunities for students to work independently, in small groups, and in large groups.	A. Designs appropriate opportunities for large group, small group, and independent student learning experiences.	A. Designs and engages students in large group, small group, and independent work activities.
B. Uses a variety of methods, including demonstrations, lectures, student initiated work, group work, questioning, and independent practice.	B. Selects from a variety of teaching methods (demonstrations, lectures, student projects, group work, independent practice) for specific instructional goals and purposes.	B. Demonstrates effective selection and use of a variety of methods to make knowledge accessible to all students.
C. Uses a variety of resources such as field trips, supplemental printed materials, manipulatives, and technology.	C. Integrates a variety of resources into instruction, including field trips, supplemental printed materials, manipulatives, and technology.	C. Demonstrates effective integration of a variety of resources and learning experiences into the curriculum.

Assessment Criteria *Benchmarks* for New Mexico Teacher Competencies for Licensure Levels I, II, and III

II. The teacher appropriately utilizes a variety of teaching methods and resources for each area taught (continued)		
D. Provides opportunities for students to apply, practice, and demonstrate knowledge and skills learned through various modalities.	D. Demonstrates understanding and appropriate application of learning styles, modalities, and intelligences theories.	D. Designs opportunities for students to apply, practice, and demonstrate knowledge and skills based on knowledge of learning modalities, style preferences, and intelligences.
E. Implements necessary modifications and adaptations in instruction and curriculum so that students with disabilities have access to the general education curriculum in the least restrictive environment.	E. Designs and implements necessary modifications and adaptations in instruction and curriculum so that students with disabilities have access to the general education curriculum in the least restrictive environment.	E. Engages with colleagues and parents to collaboratively design and implement necessary modifications and adaptations in instruction and curriculum so that students with disabilities have access to the general education curriculum in the least restrictive environment.

3. The teacher communicates with and obtains feedback from students in a manner that enhances student learning and understanding.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Explains and/or demonstrates the relevance of topics and activities.	A. Effectively explains, demonstrates or communicates the relevance of topics and activities.	A. Engages students in explaining and/or demonstrating the relevance of topics and activities.
B. Communicates to students the instructional intent, directions, or plan.	B. Consistently communicates to students the instructional intent, directions, and plans.	B. Involves students in establishing instructional direction and plans.
C. Establishes and states expectations for student performance.	C. Establishes and states expectations for student performance	C. Establishes and states expectations for student performance.
D. Clarifies actions, directions, and explanations when students do not understand.	D. Presents directions and explanations in a variety of ways to insure student understanding.	D. Presents directions and explanations in a variety of ways to insure student understanding.
E. Actively solicits communication from students about their learning.	E. Solicits communication from students about their learning for the purposes of ongoing instructional planning.	E. Engages students in the analysis and evaluation of their learning and adjusts instruction based on student feedback.
F. Communicates regularly with students about their progress.	F. Communicates regularly with students about their progress.	F. Communicates regularly with students about their progress.

Assessment Criteria *Benchmarks* for New Mexico Teacher Competencies for Licensure Levels I, II, and III

4. The teacher comprehends the principles of student growth, development and learning, and applies them appropriately.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Instructs students in the use of cognitive thinking skills such as critical thinking, problem-solving, divergent thinking, inquiry, and decision-making.	A. Consistently integrates the use of cognitive thinking skills such as critical thinking, problem-solving, divergent thinking, inquiry, and decision-making into instruction.	A. Consistently integrates the use of cognitive thinking skills such as critical thinking, problem-solving, divergent thinking, inquiry, and decision-making into instruction.
B. Uses teaching techniques that address student learning levels, rates, and styles.	B. Adapts teaching techniques to accommodate a range of student learning levels, rates, styles and special needs.	B. Selects the most effective teaching techniques to address a variety of student learning levels, rates, styles and needs as well as diverse interests and backgrounds.
C. Uses materials and media that address student learning levels, rates, and styles.	C. Adapts materials and media to address a range of student learning levels, rates, styles and special needs.	C. Selects the most effective materials and media to address a variety of student learning levels, rates, styles and needs.
D. Uses resources such as community service agencies, school personnel, and parents to meet student learning levels, rates and styles.	D. Selects from a variety of community service agencies, specialized school personnel, and parents to address different learning levels, rates, styles, and needs.	D. Integrates community resources, service agencies, other school personnel, parents, and community members into the curriculum.

5. The teacher effectively utilizes student assessment techniques and procedures.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Uses a variety of assessment tools and strategies.	A. Selects appropriate assessment tools and strategies for specific learning outcomes.	A. Designs and uses multiple methods of measuring student understanding and growth.
B. Uses information gained from ongoing assessment for remediation and instructional planning.	B. Uses formative and summative assessment for remediation and instructional planning.	B. Integrates assessment data from multiple sources into instructional planning and improvement.
C. Maintains documentation of student progress.	C. Maintains documentation of student progress.	C. Maintains documentation of student progress.
D. Communicates student progress to students and families in a timely manner.	D. Consistently maintains communication with students and families about student progress.	D. Develops a two-way system of communicating with students and families about student progress.

Assessment Criteria *Benchmarks* for New Mexico Teacher Competencies for Licensure Levels I, II, and III

6. The teacher manages the educational setting in a manner that promotes positive student behavior and a safe and healthy environment.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Serves as a model for constructive behavior patterns.	A. Identifies, explains, and models constructive behavior patterns.	A. Integrates the teaching of constructive, pro-social behaviors into regular instruction.
B. Executes routine tasks effectively and efficiently.	B. Establishes and teaches effective and efficient routines.	B. Establishes and teaches effective and efficient routines.
C. Establishes and states expectations for student behavior.	C. Establishes and reinforces expectations for student behaviors that promote citizenship in a classroom community.	C. Engages students in establishing expectations for building a learning community in the classroom.
D. Handles transitions effectively.	D. Maintains smoothness and momentum during classroom transitions.	D. Maintains smoothness and momentum during instructional transitions.
E. Has materials and media ready for student use.	E. Prepares and arranges material in advance for easy student accessibility.	E. Establishes an environment where materials and media are available and ready for student use.
F. Minimizes distractions and interruptions.	F. Minimizes distractions and interruptions.	F. Minimizes distractions and interruptions.
G. Manages student behavior effectively and appropriately.	G. Monitors and directs student behavior effectively and appropriately.	G. Develops a classroom management system that promotes acceptable and appropriate student behavior.
H. Identifies hazards, assesses risks, and takes appropriate action.	H. Identifies hazards, assesses risks, and takes appropriate action.	H. Identifies hazards, assesses risks and takes appropriate action.

Assessment Criteria *Benchmarks* for New Mexico Teacher Competencies for Licensure Levels I, II, and III

7. The teacher recognizes student diversity and creates an atmosphere conducive to the promotion of positive student involvement and self-concept.

Provisional Teacher - LEVEL I	Professional Teacher - LEVEL II	Master Teacher - LEVEL III
A. Demonstrates sensitivity and responsiveness to the personal ideas, learning needs, interests, and feelings of students with disabilities and/or from culturally and linguistically diverse backgrounds (e.g., Native Americans, Hispanic Americans, African Americans, Asian Americans, as well as other recent immigrant groups).	A. Acknowledges and validates the ideas, learning needs, interests, and feelings of students with disabilities and/or from culturally and linguistically diverse backgrounds (e.g., Native Americans, Hispanic Americans, African Americans, Asian Americans, as well as other recent immigrant groups.	A. Adjusts practice based on observation and knowledge of students with disabilities and/or from culturally and linguistically diverse groups (e.g., Native Americans, Hispanic Americans, African Americans, Asian Americans, as well other recent immigrant groups).
B. Acknowledges student performance and achievement.	B. Consistently recognizes student performance and achievements.	B. Creates curriculum designs that include student performance and acknowledgement of achievement.
C. Acknowledges that every student can learn.	C. Understands how students differ in their approaches to learning and adjusts instruction to meet diverse needs.	C. Demonstrates an awareness of the influences of context, disability, language, and culture on student learning.
D. Provides opportunities for each student to succeed and understands how students differ in their approaches to learning based on diverse cultural and linguistic backgrounds and exceptionalities.	D. Designs opportunities for each student to succeed, based on individual learning needs.	D. Provides accommodations and interventions that allow each student to succeed based on individual learning needs.
E. Provides students with opportunities for active involvement and creativity.	E. Designs specific activities that require active involvement and creativity.	E. Engages students in learning experiences that promote creativity, critical and divergent thinking.
F. Provides opportunities for students to be responsible for their behavior and learning.	F. Designs opportunities that require and reinforce student responsibility for learning.	F. Designs opportunities that require and reinforce student responsibility for learning.
G. Promotes positive student/teacher relationships.	G. Develops students' self-esteem, motivation, character, and sense of civic responsibility.	G. Fosters the development of respect for individual, cultural, linguistic, disability, and religious differences.
H. Encourages high student expectations.	H. Establishes and communicates high expectations for all students.	H. Engages students in setting high standards for performance.
I. Demonstrates an awareness and respect for each student's background, experience, learning ability, language, and culture.	I. Demonstrates knowledge of different student backgrounds, experiences, learning abilities, languages, and cultures and incorporates this knowledge into curricular decisions and instructional methodology.	I. Treats all students equitably, recognizing and planning for individual differences in cultures, languages, learning abilities, backgrounds, and experiences.

Assessment Criteria *Benchmarks* for New Mexico Teacher Competencies for Licensure Levels I, II, and III

8. The teacher demonstrates a willingness to examine and implement change, as appropriate.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Seeks out information on methodology, research and current trends in education to enhance and improve the quality of learning.	A. Seeks out information on methodology, research and current trends in education to enhance and improve the quality of learning.	A. Demonstrates the ability to reason, take multiple perspectives, be creative, and take reasoned risks to improve teaching.
B. Implements a variety of strategies to enhance learning.	B. Demonstrates knowledge of best practices that enhance learning.	B. Collaborates with colleagues in the research and design of improved instructional strategies
C. Recognizes that change entails risk and modifications may be needed.	C. Participates in instructional improvement and school reform initiatives.	C. Assumes a leadership role in the study and implementation of instructional improvement and school reform initiatives.

9. The teacher works productively with colleagues, parents and community members.		
Provisional Teacher - LEVEL I	**Professional Teacher - LEVEL II**	**Master Teacher - LEVEL III**
A. Collaborates with colleagues.	A. Actively promotes collegial relations with other school personnel.	A. Serves as a role model for collaborative working relations across the profession.
B. Communicates with parents on a regular basis.	B. Provides a system for interactive communication between teacher and parents.	B. Demonstrates knowledge of specific school, family, and community resources that can support student learning.
C. Uses conflict resolution strategies when necessary.	C. Uses conflict resolution strategies as appropriate.	C. Assists colleagues in the use of conflict resolution strategies.
D. Involves parents and community in the learning environment.	D. Promotes active roles for parents and community members in student learning.	D. Engages parents and community members productively in the work of the school.
E. Communicates in a professional manner with colleagues, parents, and community members regarding educational matters.	E. Communicates in a professional manner with colleagues, parents, and community members regarding educational matters.	E. Works collaboratively and creatively with colleagues, parents, and community members regarding educational matters.

APPENDIX C: Comparisons Among Value-Added Models

Different Value-added Model types and Associated Advantages and Disadvantages

Model Type	Advantages	Disadvantages
Percent Passing Change Model	1. Familiar to policymakers and stakeholders 2. Simple to calculate	1. Not technically a value-added model 2. Produces teacher measures that are confounded with differences between cohorts
Average Score Change Model	1. Familiar to policymakers and stakeholders 2. Simple to calculate	1. Requires vertically scaled scores across grades 2. Does not control for student characteristics that are unrelated to teacher effectiveness
Multiple Regression Model	1. Estimates teacher effectiveness as the residualized gain in student's current score after controlling for student's prior performance and demographic characteristics 2. Does not require vertically scaled scores across grades	1. Moderately complex and hard to explain to policymakers and stakeholders 2. Does not account for grouping effects
Hierarchical Linear Regression Model	1. Accounts for grouping of students within teachers 2. Estimates teacher effectiveness as the residualized gain in student's current score after controlling for student's prior performance and demographic characteristics 3. Does not require vertically scaled scores across grades	1. Highly complex and hard to explain to policymakers and stakeholders
Layered Mixed Effects Model	1. Apportions credit for student score gains to individual teachers 2. Does not require vertically scaled scores across grades	1. Highly complex and hard to explain to policymakers and stakeholders 2. Has stringent data requirements

Source: Wei, Hembry, Murphy & McBride, 2012

APPENDIX D: Value-added Model Methodology

VAM estimates were calculated using a dataset with all teachers with at least 10 tested students for five school years. Student test scores were normalized within year and test, and calculations were made using the resulting Z-scores. The estimates were the Best Unbiased Linear Predictors of score gains after the mixed-effects model was run. Two types of models were run: one type included the teaching context (student demographic model), and one type did not (test score only model). Models with one and two years of prior student scores were run. Students who did not have either prior test score or demographic information were excluded from the models where those covariates were required.

Teacher and school-level contextual covariates were the averages of their individual level variables.

The reduced form equations associated with the models are as follows:

Two year student demographic model: $Z_{i,t} = Z_{i,t-1}\alpha_g + Z_{i,t-2}\delta_g + \boldsymbol{X_{i,t}\beta} + \boldsymbol{\bar{X}_{c,t}\gamma_c} + \boldsymbol{\bar{X}_{s,t}\gamma_s} + (\mu_{t,t} + \varepsilon_{i,t})$

Two year test score only model: $Z_{i,t} = Z_{i,t-1}\alpha_g + Z_{i,t-2}\delta_g + (\mu_{t,t} + \varepsilon_{i,t})$

One year student demographic model: $Z_{i,t} = Z_{i,t-1}\alpha_g + \boldsymbol{X_{i,t}\beta} + \boldsymbol{\bar{X}_{c,t}\gamma_c} + \boldsymbol{\bar{X}_{s,t}\gamma_s} + (\mu_{t,t} + \varepsilon_{i,t})$

One year test score only model: $Z_{i,t} = Z_{i,t-1}\alpha_g + (\mu_{t,t} + \varepsilon_{i,t})$

Subscripts in the models are as follows: i is the individual student, g is the grade, t is the teacher. When preceded by a comma, t is the time period, where t -1 is the prior time period, and t -2 is the time period two periods prior.

Covariates included in the contextual model at the individual level were: FRL status, gender, ethnicity, sped level/gifted, ELL status, FAY status, old/young/repeating grade.

Covariates included in the contextual model at the school and teacher level were: FRPL status, gender, ethnicity, sped level/gifted, ELL status, FAY status, old/young/repeating grade, average and standard deviation of prior math and reading scores.

The teacher effect estimate is the Best Linear Unbiased Predictor of $\mu_{t,t}$. It is calculated using a random effects specification. $\boldsymbol{X_{i,t}}$,$\boldsymbol{\bar{X}c,t}$, **and** $\boldsymbol{\bar{X}_{s,t}}$ are the student, teacher, and school-level vectors of covariate averages, respectively.

Report
to
The LEGISLATIVE FINANCE COMMITTEE

Public Education Department
Developing Early Literacy in New Mexico
July 12, 2012

Report #12-05

LEGISLATIVE FINANCE COMMITTEE

Senator John Arthur Smith, Chairman
Representative Luciano "Lucky" Varela, Vice-Chairman
Senator Sue Wilson Beffort
Senator Pete Campos
Senator Carlos R. Cisneros
Representative William "Bill" J. Gray
Senator Stuart Ingle
Representative Rhonda S. King
Representative Larry A. Larrañaga
Senator Carroll H. Leavell
Senator Mary Kay Papen
Representative Henry "Kiki" Saavedra
Representative Nick L. Salazar
Representative Edward C. Sandoval
Senator John Sapien
Representative Don L. Tripp
Representative James P. White

DIRECTOR

David Abbey

DEPUTY DIRECTOR FOR PROGRAM EVALUATION

Charles Sallee

PROGRAM EVALUATION TEAM

Jeff Canney, CGFM
Jon Courtney, Ph.D.
Valerie Crespin-Trujillo
Jack Evans
Brenda Fresquez, CICA
Pamela Galbraith
Maria D. Griego
Rachel Mercer-Smith
Matthew Pahl
Elaine Romero
Michael Weinberg, Ed.D.

State of New Mexico
LEGISLATIVE FINANCE COMMITTEE
325 Don Gaspar, Suite 101 • Santa Fe, NM 87501
Phone: (505) 986-4550 • Fax (505) 986-4545

David Abbey
Director

July 12, 2012

Ms. Hanna Skandera, Secretary-Designate
Public Education Department
Jerry Apodaca Education Building
300 Don Gaspar Avenue
Santa Fe, NM 87501

Dear Ms. Skandera:

On behalf of the Legislative Finance Committee (Committee), I am pleased to transmit the evaluation, *Developing Early Literacy in New Mexico.* The program evaluation team assessed student performance; federal, state, and spending patterns; and best practices for accelerating student achievement in literacy. The report will be presented to the Committee on July 12, 2012. Exit conferences were conducted with the Public Education Department on May 22, 2012 to discuss the contents of the report. The Committee would like a plan to address the recommendations within this report within 30 days from the date of the hearing.

I believe this report addresses issues the Committee asked us to review and hope New Mexico's public education system benefits from our efforts. We very much appreciate the cooperation and assistance we received from your staff.

Sincerely,

David Abbey

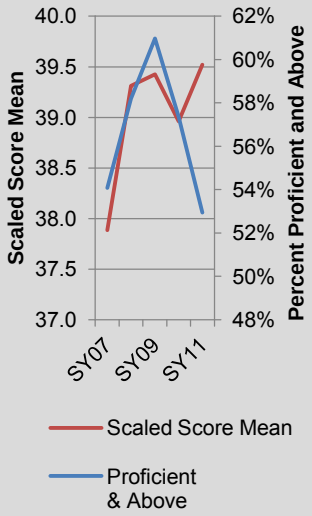
David Abbey, Director

cc: Senator John Arthur Smith, Chairman, Legislative Finance Committee
Representative Luciano "Lucky" Varela, Vice-Chairman, Legislative Finance Committee
Representative Henry "Kiki" Saavedra, Legislative Finance Committee
Representative Rick Miera, Chairman, Legislative Education Study Committee
Yolanda Berumen-Deines, Secretary, Children, Youth, and Families Department

Table of Contents

Page No.

Grade 3 Reading SBA Scaled Score and Proficiency Rate Trends



Source: LFC Analysis

Student performance on New Mexico's standards-based assessment is reported in four categories:

- ***beginning steps (0-31 points)***
- ***nearing proficient (32-39 points)***
- ***proficient (40-55 points)***
- ***advanced (56-80 points)***

Students scoring proficient are considered to be at grade-level.

In SY11, 2,446 third-graders, or 10 percent, were within 2 points of scoring proficient on the SBA.

Early reading proficiency is well-established as a strong predictor of high school graduation rates as well as future earning potential. In spite of slight improvements to scaled scores on New Mexico's standards-based assessments (SBA), third-grade reading proficiency rates continue to lag behind desired levels.

In response, New Mexico has invested heavily to improve early literacy, including full-day kindergarten, prekindergarten (PreK), and an extended-school-year program, Kindergarten-Three Plus (K-3 Plus). The Legislature has quadrupled funding for PreK since FY06 and doubled funding for K-3 Plus between FY12 and FY13. Additionally, the state is considering initiatives, such as statutorily revising its retention policy and paying for reading coaches, professional development for teachers, and additional assessments in the early grades.

This evaluation begins with an overview of how third-graders are reading as measured by the SBA, the first statewide snapshot of student performance, and describes which students are struggling. Second, the evaluation analyzes the statistical relationship between existing and proposed state initiatives and third-grade reading proficiency. Finally, the evaluation presents lessons learned in early literacy based upon visits to schools identified as "beating the odds," where students are succeeding despite being at-risk students, as well as under-performing schools.

The percentage of third-graders reading proficiently, or at grade-level, has dropped from 61 percent in 2009 to 53 percent in 2011. These declines, however, mask improvement in average scaled scores because of changes in the cut scores used to determine proficiency. Third-grade average scaled scores have steadily increased from 37.9 in 2007 to 39.5 in 2011. Gaps persist in achievement between ethnicities, but the biggest differences are strongly associated with socioeconomic status and English-language acquisition levels. These gaps highlight the importance of allocating resources to areas of greatest need and implementing research-based language acquisition models.

Investments in PreK have resulted in measureable, significant effects on third-grade reading proficiency rates. These programs generally serve more challenging populations and are improving reading skills for those participating. Similarly, K-3 Plus, the state's initiative to extend the school year at schools with the highest poverty rates, is making a difference on student performance, particularly when combined with PreK. These interventions appear to be cost-effective alternatives to retaining students, which, based on an analysis of third-graders from 2011, has a mixed relationship with reading proficiency.

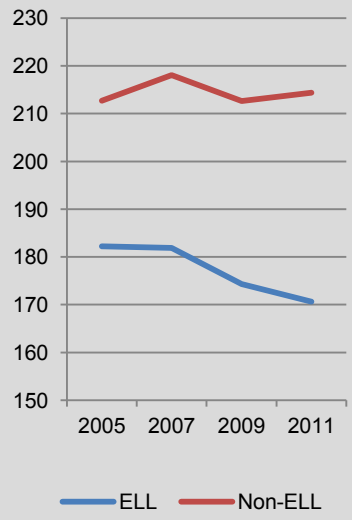
Finally, the state has an opportunity to promote outstanding teaching and leadership at the school level. High quality implementation of best practices, including using data to drive instructional decisions,

Demographics of Third-Graders, SY11
(n = 25,495)

	Percent of Total	Reading SBA Percent Proficient or Above
All Students	100%	53%
Male	50%	48%
Female	48%	58%
Hispanic	60%	48%
Caucasian	24%	70%
Native American	9%	36%
Free or Reduced-Price Lunch	72%	46%
English-language Learners	20%	33%
Students with Disabilities	15%	22%

Source: LFC Analysis

NAEP 4th Grade Reading Average Scaled Scores for ELL and non-ELL Students



Source: Nation's Report Card

systematically addressing chronic absenteeism, and aligning human and fiscal resources, correlates with higher reading proficiency rates. As this evaluation details, focus on these areas will result in meaningful reading gains in New Mexico.

KEY FINDINGS

Student performance is highly influenced by economic status, language status, and student attendance. Of a third-grade cohort of 26 thousand students from the 2011 school year (SY11), 53 percent of students were considered proficient or above in reading as measured by New Mexico's standards-based assessment (SBA). Students who are Hispanic, Native American, English-language learners (ELL), or qualified for free or reduced-price lunch (FRL) had lower proficiency rates than the overall state average.

Changes in cut scores used to determine SBA reading proficiency rates appear to mask improvements in scaled scores. New Mexico's third-grade reading proficiency rate increased to 61 percent in SY09 but has since dropped to 53 percent. In contrast, the average scaled score used to determine those proficiency rates steadily increased from 37.9 in SY07 to 39.5 in SY11. These trends in increasing scaled scores are true for all subgroups, with the exception of English-language learners, whose scaled scores have remained flat since SY07.

The largest achievement gaps in New Mexico are strongly associated with poverty and language. Nationally, New Mexico has the second highest percentage of students who qualify for FRL and the third highest percentage of English-language learners. While there are gaps in proficiency rates among ethnic subgroups, the gaps related to socioeconomic status and English-language learner status within each ethnic subgroup are even larger. On the 2011 reading SBA, for example, 14 percent of Native American FRL, ELL students were proficient compared with 64 percent of non-FRL, non-ELL Native American students.

For all subgroups, the gap between proficiency rates for FRL and non-FRL students is approximately 20 percentage points. Hispanic students have the widest range between proficiency rates for FRL and non-FRL students at 24 percentage points.

Similarly, on the National Assessment of Education Progress (NAEP), ELL students in New Mexico scored 42 points less than non-ELL students on the 2011 fourth-grade reading test. This gap grew from 2005, when ELL students scored 31 points less than non-ELL students.

At-risk students are more likely to have high rates of absenteeism, which negatively impacts achievement levels. Based on third-grade reading scaled scores from 2011, student attendance is directly related to reading proficiency. On average, each 1 percent increase in attendance, two days of school, equated to a 0.43 point scaled score increase. Based on this

Achievement Gap Between Percent Proficient FRL and Non-FRL Third-Graders, SY11

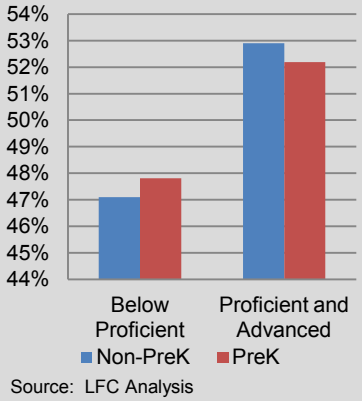
	FRL	Non-FRL	Difference
Hispanic	44%	68%	24%
Caucasian	61%	80%	19%
Native American	33%	55%	22%

Source: LFC Analysis

Recognizing the importance of attendance, the principal at Griegos Elementary in Albuquerque regularly makes calls and home visits to encourage students to come to school.

According to the National Center for Children in Poverty, preschool-aged children growing up in low-income households score 60 percent below children in the highest socioeconomic group on cognitive tests and know only a third of the words of their middle-income peers.

PreK vs. Non-PreK SBA Reading Proficiency Rates, SY11



Source: LFC Analysis

relationship, improving a student's attendance rate by 2.3 percent, roughly four school days per school year, corresponds with a one-point increase on the reading SBA scaled score.

Performance is lowest for chronically absent students. For the 1,180 students who attended school less than 90 percent of the time between first and third grade, the average SBA scaled score was 36.9 and only 43 percent were proficient. In contrast, the 15 thousand students present 95 percent or more of the school year, 56 percent were proficient with an average reading scaled score of 40.5.

Several groups of students are over-represented in this high-absence category, including Native Americans, English-language learners, and students with disabilities.

Early education improves performance, but lack of coordinated resources and inconsistent quality limits success. Children from low-income homes often start out behind and must learn more than a year's worth of academic content each school year to catch up to their more affluent peers. In 2008, PreK students scored in the 23rd percentile nationally for receptive vocabulary, a key indicator of school success.

New Mexico, however, lacks a common assessment of kindergarten readiness, making it difficult to compare programs and clearly communicate expectations to parents prior to kindergarten.

New Mexico, along with the federal government, spends more than $117 million on early learning programs like PreK, Head Start, and special education prekindergarten that help narrow, but not fully close, achievement gaps. For FY13, the state increased its FY12 PreK funding levels by 33 percent, appropriating $19 million to the Public Education Department (PED) and the Children, Youth and Families Department (CYFD). In addition to New Mexico PreK, in FY11, Head Start, a federally funded early childhood program not administered by PED or most school districts, served 5,400 four and five-year olds in New Mexico at a cost of $57 million.

PreK boosts student performance, including third-grade reading scores. In SY11, third-graders who attended New Mexico PreK were proficient at nearly identical rates, 52 percent, as the overall population of New Mexico third-graders, 53 percent. This is remarkable because these PreK programs serve higher percentages of Hispanic, Native American, ELL, and FRL students than the overall population of third-graders.

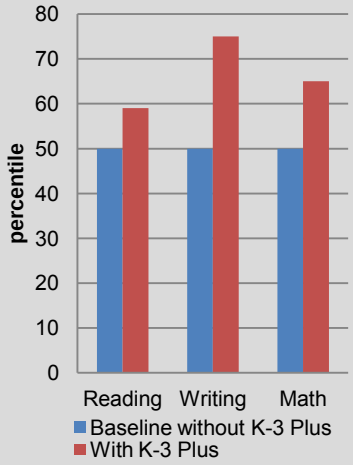
Of the 18,250 FRL third-graders in 2011, 7 percent, or 1,335 students, attended PreK. At FY13 funding levels, it is estimated that 53 percent of New Mexico FRL four-year olds receive PreK or Head Start.

PreK Enrolls Higher Percentages of High-needs Students

	CYFD PreK (778 students)	PED PreK (824 students)
Hispanic	73%	57%
Native American	7%	31%
FRL	76%	90%
ELL	15%	32%

Source: LFC Analysis

Effect of K-3 Plus on SBA Percentile Rank, SY10



Source: Utah State University

Utah State University is using a control group to evaluate the effects of K-3 Plus with results expected later this year.

PreK students outscored similar non-PreK students on the reading portion of the SBA when controlling for demographic variables. Compared with a student group made up of similar proportions of Hispanic, Caucasian, Native American, ELL, FRL, and special education students, those who participate in PreK are estimated to earn an additional 1.2 scaled score points on the third-grade reading SBA than non-participants. For third-graders who attended PreK in SY07, the estimated effect of CYFD PreK was 0.4 scaled score points; for PED PreK, the estimated effect was 1.8 scaled score points.

As measured by the National Institute for Early Education Research (NIEER) as well as New Mexico's Early Learning Outcomes Assessment (PreK assessment), New Mexico PreK produces consistent benefits for children. Based on the SY11 results of the PreK assessment overseen by the University of New Mexico, 90 percent of students are progressing across the seven measured domains, but it is unclear how many of these students are considered ready for kindergarten.

New Mexico's K-3 Plus program is making positive differences for the neediest students. K-3 Plus was developed in 2007 to narrow the achievement gap by extending the school year by a minimum of 25 days at schools with at least 85 percent of students qualifying for free or reduced-price lunch. In FY13, New Mexico appropriated $11 million to the K-3 Plus program, more than double the FY12 appropriation of $5.3 million. More students who enroll in K-3 Plus are poor, Hispanic, Native American, or English-language learners than the overall third-grade population.

A 2011 evaluation of K-3 Plus conducted by Utah State University found positive effects on third-grade reading, writing, and math SBA performance and estimated the benefits from reduced grade retention and remediation services offset all of the costs to fund K-3 Plus.

Additionally, students who receive two years of K-3 Plus outperformed students who attended one year of K-3 Plus. Controlling for student demographics, students who attended K-3 Plus a second year are estimated to score 0.8 scaled score points higher than students who attended K-3 Plus for one year.

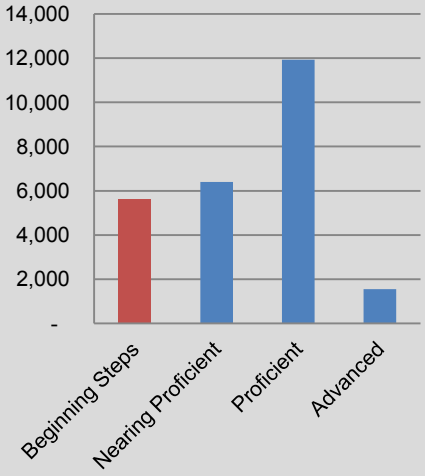
Few at-risk students have access to a full continuum of early childhood education programs, despite the need for extra learning time. In SY12, only 13 schools in New Mexico offered PreK and K-3 Plus. Of 25 thousand third-graders in New Mexico in 2011, only 81 attended PreK and two years of K-3 Plus.

Controlling for ethnicity, ELL status, special education status, and FRL status, students who received all three programs are estimated to score between 1.5 and 3.8 scaled score points higher than students who did not receive these programs.

Retained students are two to 11 times more likely to drop out of school.

In 2010, the unemployment rate for those without a high school diploma was 5 percent higher than for workers with only a high school diploma.

Number of 3rd Grade Students by SBA Reading Level, SY11



Source: LFC Analysis

Approximately 8 percent of 2011 third-graders, or 1,923 students, would have been eligible for retention under recently considered legislation.

Nine percent of the 219 third-graders retained in 2010 were proficient in reading and some dropped a reading level in their second year of third-grade.

Better attention is needed on regular performance reporting and adherence to quality of K-3 Plus program implementation. An evaluation of K-3 Plus in Albuquerque found programs were more successful if K-3 Plus teachers were paired with the same students they were teaching in the upcoming school year. Similarly, a RAND Corporation study identified maximizing quality, enrollment, and attendance as critical elements to achieving benefits from summer learning programs. While the K-3 Plus application requires plans for dedicated reading and math blocks as well as intervention services for students in the lowest quartile, the PED's ability to ensure proper implementation is limited.

Early literacy initiatives, such as mandatory retention policies and Reading First, have produced mixed results. Generally, states with mandatory retention policies focus on third-graders reading below proficient and provide remediation and intervention to increase the number of proficient third-graders. The additional support students receive includes summer reading camps and tutoring during the school year, making it difficult to isolate the causes of changing student achievement.

The Legislature has recently considered revisions to New Mexico's remediation and retention laws. Current law allows parents of students in grades one through eight to refuse to allow their child to be retained for one year. Schools are then responsible for developing and implementing academic improvement plans, and if the student does not make sufficient progress, that student may be retained.

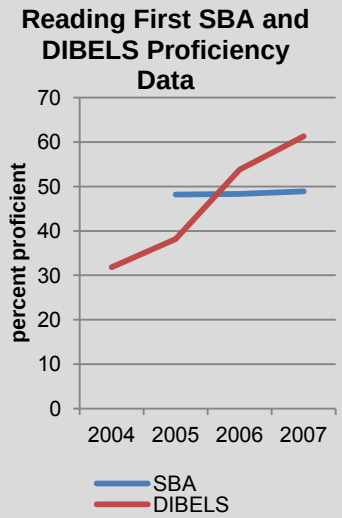
Recently considered legislation would require retention of third-grade students scoring at beginning steps, the lowest proficiency level on the SBA. Of the 25,495 third-graders with valid SBA reading scores in 2011, 5,628, or 22 percent, met this criteria; 6,392 students, or 25 percent, were nearing proficient, but would not be mandatorily retained. Consistent with similar legislation in other states, students would be exempt from retention for proficient scores on an alternate assessment, completion of a portfolio, ELL status, special education status, or previous retention.

On average, providing an additional school year costs the state $7,000, the same amount as providing PreK and four years of K-3 Plus per student. For the 1,923 third-graders eligible for retention in SY11, this would have cost the state $13.5 million, compounding with each cohort subject to a mandatory third-grade retention policy.

In New Mexico, almost 10 percent of the third-grade class in SY11 had been retained between kindergarten and third-grade. The greatest number of retained students, 936, were kindergarteners, decreasing steadily to 232 retained in the third-grade. Of those retained students, 30 percent qualified for special education services by the time they were third-graders. Of all students retained in kindergarten to third grade, only 29 percent were proficient as third-graders in 2011.

Of the retained third-graders who scored at beginning steps in SY10, only 12 percent were proficient by the end of their second year of third-grade. Of the 100 third-graders who scored at beginning steps and were retained in SY10, 25 stayed at beginning steps in SY11, 63 moved up one level to

While DIBELS scores rose during the Reading First initiatives, state SBA scores remained flat.



Source: Reading First Annual Performance Report

Of the six high-performing schools the LFC evaluation team visited, Dolores Gonzalez, Washington Avenue, Griegos, Mesilla Park, Jaramillo, and Newcomb Elementary, each had principals with tenures of 10 years or more. In contrast, of the six low-performing schools visited, each had at least three and as many as five principals in the last 10 years.

Over the last five years, the number of school leaders prepared by New Mexico's five higher education institutions has decreased by 38 percent.

Level III teachers earn a minimum of $278 per day, while elementary principals earn $273 per day.

nearing proficient, and 12 moved to proficient. Although many students increased their reading scores in their second year of third-grade, overall, only 29 percent moved up to become proficient in reading after repeating the third-grade.

During Reading First, a federally funded program similar to the currently proposed New Mexico Reads to Lead initiative, third-grade reading proficiency rates did not improve. From 2002 to 2008, New Mexico received $63 million in federal Reading First funds. In 2007, these funds served 19 thousand students in 100 schools across 39 districts.

Similar to the Reads to Lead initiative, Reading First required states and participating school districts to adopt scientifically based reading programs, provide professional development, and track students' reading progress using valid and reliable assessments. Almost all states required Reading First schools to have a reading coach to support teachers and principals with instruction, assessments, and data interpretation.

As measured by the dynamic indicators of basic early learning skills (DIBELS), the percentage of third-graders participating in Reading First reading fluently increased from 23 percent in 2004 to 61 percent in 2007. Over this same time, however, SBA reading proficiency rates remained flat at 48 percent, highlighting differences between the DIBELS, which measures fluency, and the SBA, which measures comprehension.

State, district, and school-level management practices can help schools to marginally beat the odds. Quality teaching is the most important school-based influence on student achievement. In 2009, the National Council on Teacher Quality evaluated eight New Mexico teacher preparation programs' and found low admissions standards; lack of focus on the science of teaching reading, including poor reading textbook quality; inadequate math preparation, including poor math textbook quality; and a weak exit assessment.

School leadership also has an indirect effect on student learning. Numerous researchers have established that strong leadership has a statistically significant influence on student achievement, accounting for up to 25 percent of the total school effect (Marzano, 2000). Based on an LFC survey, schools with less principal turnover average higher scaled scores than those with higher principal turnover. Additionally, schools with three principals in the past 10 years have nearly double the rate of teacher turnover as schools with only one principal.

Elements within the current administrator licensure requirements and minimum salary structure act as obstacles to the supply of qualified school leaders in New Mexico. Becoming a principal in New Mexico, for example, requires a minimum of six years of teaching experience, while neighboring states require only two to three years.

High quality implementation of best practices impacts student growth as measured by the SBA. Based on analysis of an LFC survey, for example, the use of the Measures of Academic Progress as a short-cycle assessment correlates with student performance on the SBA. Similarly, schools that

At Jaramillo Community School in Belen, everyone from the principal to the literacy coach to the teachers uses DIBELS data to regularly monitor student progress and drive instructional decisions.

Gus Benakis, former principal of Harrison Schmitt Elementary in Silver City, identified 10 "Characteristics of Success":

1. ***Raise the expectations, clarify the focus;***
2. ***Communicate (Listen!);***
3. ***Be visible (especially the principal);***
4. ***Collaborate K through 5;***
5. ***Identify/ assess student needs early;***
6. ***Retain early (K, 1);***
7. ***Align curriculum (especially in weak areas);***
8. ***Provide professional development;***
9. ***Embrace challenges and acknowledge success; and***
10. ***Take away the excuses and provide necessary resources.***

indicate their reading coaches spend time analyzing data were found to have better student outcomes. Finally, regularly using the DIBELS to monitor student progress was also found to correlate with higher SBA scores.

KEY RECOMMENDATIONS

The Legislature should require districts to annually submit performance-based budgets to the PED, consistent with the General Appropriations Act.

The Legislature should require a minimum of three years of teaching, or level II licensure, to obtain an administrative license.

The PED should:

Annually report its process for determining standards-based assessment cut scores and the relative impact on proficiency rates to the Legislature;

Evaluate the impact of bilingual models on the performance of New Mexico's English-language learners;

Require districts to report data on student attendance and identify strategies for improvement as part of its annual performance-based budgeting process;

Raise attendance criteria in school grading to encourage schools to improve attendance rates;

Track student enrollment in PreK, Head Start, or other pre-school programs in the student teacher accountability reporting system (STARS); also, collect New Mexico PreK assessment data in STARS;

Coordinate allocation of PreK and K-3 Plus resources to increase the number of students who receive the full benefit of both programs;

Increase oversight and accountability of K-3 Plus to improve consistency and quality of implementation;

Require districts to report data on principal and teacher turnover and identify strategies for improvement;

Adopt statewide, short-cycle assessments in grades K-3 that align to the common core standards, measure growth of all students at least three times per year, can be used more frequently to monitor the progress of higher-need students, and allow comparisons with other states.

The PED and the CYFD should consider alternative PreK assessments based on: cost effectiveness, time required for administration, alignment with the common core kindergarten standards, and ability for comparisons with other states.

Academically, success in third-grade reading is critical. A 2011 Annie E. Casey Foundation study found children who read proficiently in third grade are four times more likely to graduate from high school than non-proficient third-graders. Additionally, students who graduate from high school earn nearly 25 percent more than non-high school graduates, and over half of high school graduates are employed full-time compared with 38 percent of non-high school graduates.

New Mexico has initiated a number of programs to improve early literacy, including full-day kindergarten, prekindergarten (PreK) and Kindergarten-Three Plus (K-3 Plus). Federal programs, such as Reading First and the 21st Century Community Learning Centers, have also focused on reading in kindergarten through third grade and extended learning time for high-needs students. The costs for these programs are considerably lower per student than the estimated cost to the state of $250 thousand per high school dropout for public assistance programs and efforts to offset the dropout's reduced contribution to society

In the early 1990s, changes in formula factors resulted in greater funds allocated to kindergarten through third grade, but because allocations to districts are non-categorical, this does not necessarily mean more was spent in those grades. Legislation enacted in 1993 changed formula cost factors to adequately fund maximum average class loads for elementary school classrooms. The law capped first, second, and third-grade classrooms at an average of 22 students and kindergarten classrooms at 20 students. Formula factors also increased for fourth, fifth, and sixth grades.

Students in grades kindergarten through third grade generated $488 million from the general fund through New Mexico's formula in FY11, 21 percent of all public school funding through the formula. Students in grades one to three were funded at $4,300 per student, more than the per-student funding for grades four through six, but less than the $4,500 per student in grades seven through 12.

Table 1. FY11 Per-Student Funding by Grade Level

Grade	Per Student Funding
K	$5,164
1st	$4,303
2nd and 3rd	$4,231
4th through 6th	$3,747
7th through 12th	$4,482

Source: PED

This evaluation analyzes the relationship between state initiatives and third-grade reading proficiency as measured by New Mexico's standards-based assessments (SBA), the first statewide snapshot of student performance in reading and math.

Given the potential impact of early literacy investments on future student success, the Legislative Finance Committee is evaluating educational policy and programs designed to increase statewide early reading proficiency to identify best practices and ensure efficient and effective use of public resources (see **Appendix A** for project information and **Appendix B** for the PED Performance Report Card).

STUDENT PERFORMANCE IS HIGHLY INFLUENCED BY ECONOMIC AND LANGUAGE STATUS AND STUDENT ATTENDANCE

Each year, little more than half of New Mexico's 25 thousand third-graders demonstrate reading success. Of the 2011 third-grade cohort (SY11), 53 percent of students were considered proficient or above in reading as measured by New Mexico's standards-based assessment (SBA). Students who are Hispanic, Native American, English-language learners (ELL), or qualified for free or reduced-price lunch (FRL) had lower proficiency rates than the overall state average. See **Appendix C** for proficiency rates by school.

The New Mexico Standards-Based Assessment (SBA)
In SY11, the 3rd grade reading portion of the SBA consisted of three sessions totaling approximately 160 minutes and was made up of 33 multiple choice items, five short answer items, and three open-ended response items, for a total of 65 possible points. **Appendix D** provides a sample open-ended response item. These point totals are then converted to scaled scores and are reported in four categories: beginning steps (0-31 points), nearing proficient (32-39 points), proficient (40-55 points), and advanced (56-80 points).

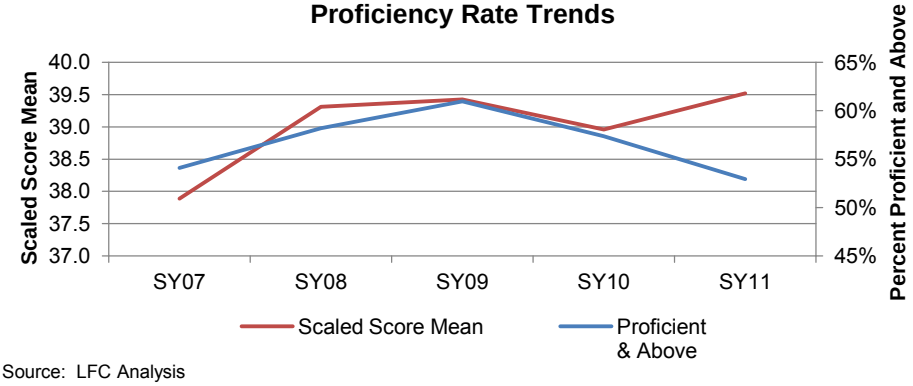
Table 2. Demographics of Third-Graders, SY11

	Number	Percent of Total	Reading SBA Percent Proficient or Above
All Students	25,495	100%	53%
Male	12,618	50%	48%
Female	12,182	48%	58%
Hispanic	15,302	60%	48%
Caucasian	6,105	24%	70%
Native American	2,330	9%	36%
Other	1,063	4%	57%
FRL	18,250	72%	46%
ELL	5,204	20%	33%
Students with Disabilities	3,914	15%	22%

Source: LFC analysis of PED data

Changes in cut scores used to determine SBA reading proficiency rates appear to mask improvements in scaled scores. Reading proficiency rates increased to 61 percent in SY09 but have since dropped to 53 percent. In contrast, the scaled scores used to determine those proficiency rates steadily increased from 37.9 in SY07 to 39.5 in SY11. These trends in increasing scaled scores are true for all sub-groups, with the exception of English-language learners, whose scaled scores have remained flat since SY07.

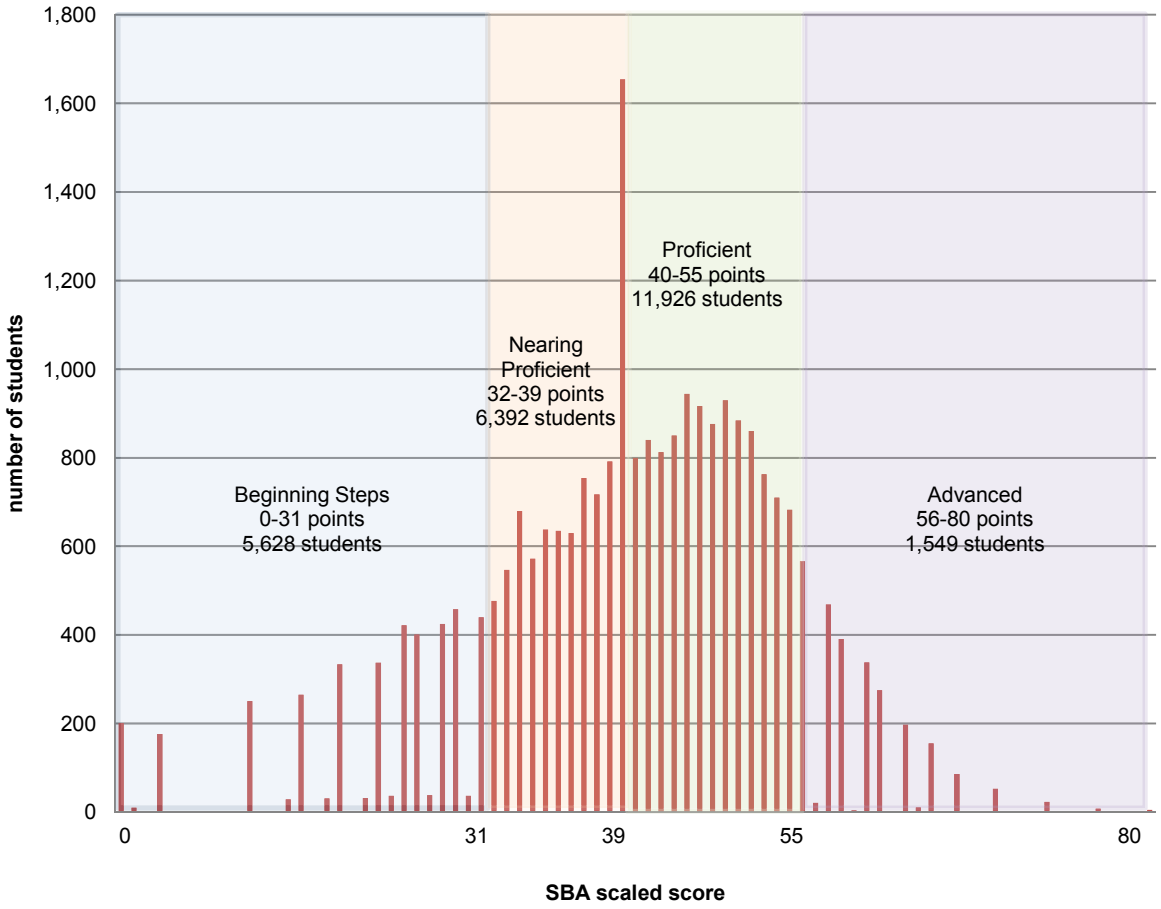
Chart 1. Grade Three Reading SBA Scaled Score and Proficiency Rate Trends



Source: LFC Analysis

In SY11, 2,446 third-graders, or 10 percent, were within two points of scoring proficient on the SBA. Similarly, improving scaled scores by two points would move an additional 5 percent of students into the nearing proficient category. Minimal improvements on the SBA would significantly alter New Mexico's achievement levels.

Chart 2. Frequency of Grade Three Reading SBA Scaled Scores, FY11

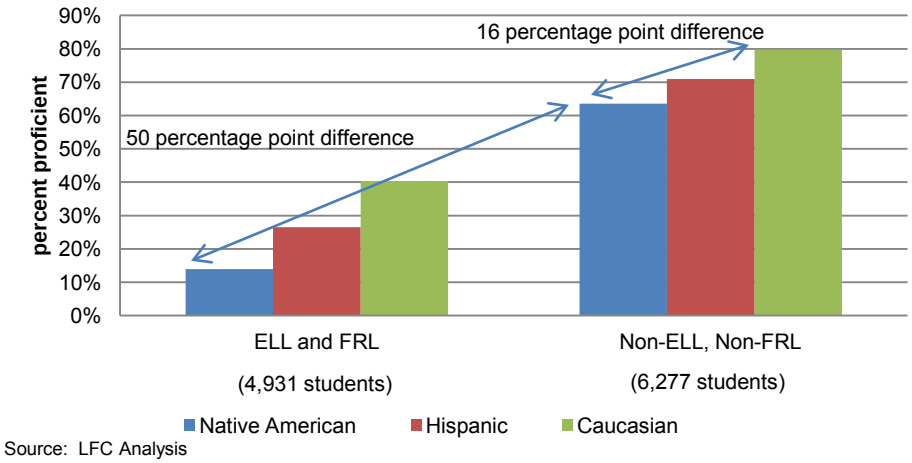


Source: LFC Analysis

The largest achievement gaps in New Mexico are strongly associated with poverty and language. Research has established a strong relationship between student performance and both economic and language status (Lee & Burkam, 2002). Nationally, New Mexico has the second highest percentage of students who qualify for free or reduced-price lunch (FRL) and the third highest percentage of English-language learners (ELL). The distribution of FRL students can be seen in **Appendix E**.

While there are gaps in proficiency rates among ethnic subgroups, the gaps related to socioeconomic status and English-language learner status within each ethnic subgroup are even larger. On the 2011 reading SBA, for example, 14 percent of Native American FRL, ELL students were proficient compared with 64 percent of non-FRL, non-ELL Native American students. This same gap of roughly 50 percentage points is true for each ethnic subgroup and is consistent with the LFC's analysis of SBA scores from 2004 to 2008 presented in the "Three-Tiered Licensure System and the Achievement Gap" in 2009.

Chart 3. FRL and ELL Third-Grade Reading Achievement Gap, FY11



Source: LFC Analysis

Similarly, for all subgroups, the gap between proficiency rates for FRL and non-FRL students is approximately 20 percentage points.

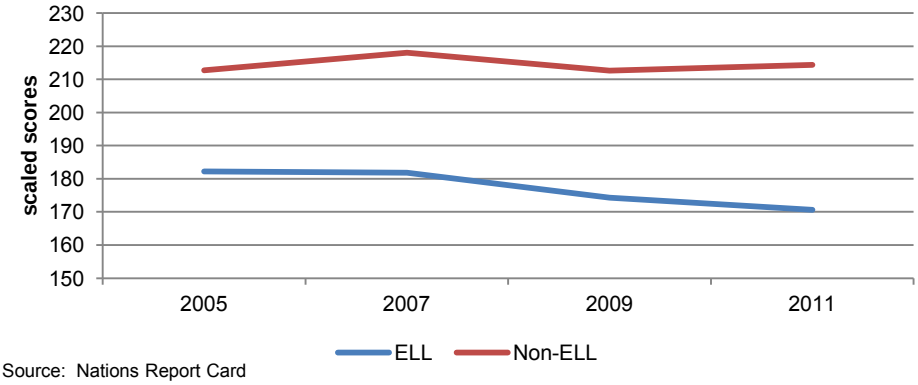
Table 3. Achievement Gap Between Percent Proficient FRL and Non-FRL Third-Graders, SY11

	FRL	Non-FRL	Difference
Hispanic	44%	68%	24%
Caucasian	61%	80%	19%
Native American	33%	55%	22%
Other	48%	76%	28%

Source: LFC Analysis

On the National Assessment of Education Progress (NAEP), also known as the Nation's Report Card, ELL students in New Mexico scored 42 points less than non-ELL students on the 2011 fourth-grade reading test. This gap grew from 2005, when ELL students scored 31 points less than non-ELL students.

Chart 4. NAEP Fourth-Grade Reading Average Scaled Scores for ELL and non-ELL Students



Source: Nations Report Card

The percentage of fourth-grade ELL students scoring proficient and above on the NAEP has also declined. In 2003, 9 percent of New Mexico's ELL population was proficient or above on the fourth-grade reading NAEP; in 2011, 2 percent of ELL students scored proficient and above. National trends showed 8 percent of ELL students were proficient on NAEP reading in 2003 and 2011 (**Appendix F**).

Students scoring below proficient are concentrated by district and school. Of the 25,495 valid third-grade SBA scores in 2011, 12,020 scored beginning steps or nearing proficient. More than half of those students, 6,886, came from seven school districts. These same districts educate 56 percent of New Mexico's third-graders, suggesting a relatively even distribution of non-proficient third-grade readers across the state.

Table 4. Third-grade Students Below Proficient by District, SY11

District	Number of students below proficient	Percent of total students below proficient	District's Percent of Third Graders in NM
Albuquerque	3,205	27%	29%
Las Cruces	965	8%	7%
Gallup McKinley	633	5%	3%
Santa Fe	590	5%	5%
Gadsden	584	5%	4%
Rio Rancho	479	4%	5%
Farmington	430	4%	3%
Total	**6,886**	**57%**	**57%**

Source: LFC Analysis

While those students are spread across 425 schools, 9 percent, or 1,030 students came from 10 elementary schools. Expanding this list to the top 30 schools covers 20 percent of the states' students reading below proficient.

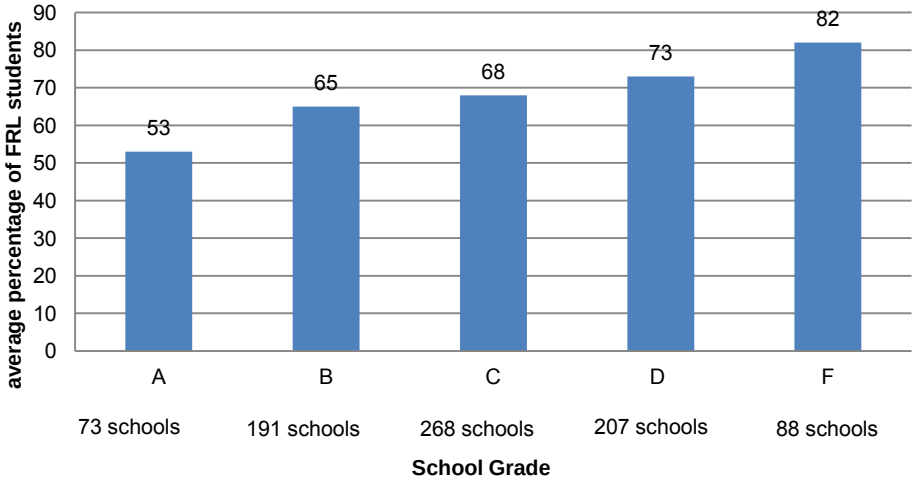
Table 5. Students Below Proficient by School, SY11

School (District)	Number of third-graders below proficient	Total number of third-graders	Percent of third-graders below proficient
Edward Gonzales (APS)	130	207	63%
Central (Bloomfield)	119	214	56%
Navajo (Gallup McKinley)	107	142	75%
MaryAnn Binford (APS)	106	161	66%
Jefferson (Lovington)	105	220	48%
Valencia (Portales)	103	238	43%
Carlos Rey (APS)	94	151	62%
Painted Sky (APS)	93	182	51%
Sunrise (Las Cruces)	91	150	61%
W.D. Carroll (Bernalillo)	80	140	57%
Total	**1,028**	**1,805**	

Source: LFC analysis

The PED's assignment of school grades favored schools with fewer students from low-income families. At schools earning an A, an average of 53 percent of students were FRL compared with 82 percent at schools earning an F.

Chart 5. Average Percentage of Free or Reduced-Price Lunch Students and All School Grades, FY11



Source: LFC analysis of PED data

<u>At-risk students are more likely to have high absenteeism, which negatively impacts achievement levels.</u> Based on third-grade reading scaled scores from 2011, student attendance is directly related to reading proficiency. The average three-year attendance rate for third-graders scoring beginning steps on the SBA was 95.4 percent, while the attendance rate for students scoring at the highest level, advanced, was 96.8 percent. A 1 percent difference is approximately two days per school year. The average rate of attendance for 21,639 third-graders between SY09 and SY11 was 96 percent.

Table 6. Three-Year Attendance Rates of SY11 Third-Graders

	Number of students	Average Attendance Rate
Beginning Steps	4,618	95.4%
Nearing Proficient	5,405	95.9%
Proficient	10,278	96.3%
Advanced	1,338	96.8%
Total	21,639	96.0%

Source: LFC Analysis

On average, each 1 percent increase in attendance equated to a 0.43 point scaled score increase. Based on this relationship, improving a student's attendance rate by 2.3 percent, roughly four school days per school year, corresponds with a one-point increase on the reading SBA scaled score. As seen in Chart 2, 6 percent of New Mexico's third-graders were within one point of proficient and 12 percent were within three points.

Performance is lowest for chronically absent students. For the 1,180 students who attended school less than 90 percent of the time between first and third grades, the average SBA scaled score was 36.9, with 40 being proficient, and only 43 percent were proficient. In contrast, of the 15 thousand students present 95 percent or more of the school year, 56 percent were proficient with an average reading scaled score of 40.5. Over-represented students with high-absences include Native Americans, English-language learners, and special education.

Student ethnicity and poverty status directly correlate with attendance rates and SBA reading scaled scores. Native American students tended to have the lowest attendance rates, 1.4 percentage points below the state average, and reading SBA scaled scores 4.3 points below the state average of 39.5.

Table 7. Relationship between Attendance and SBA Performance for Subgroups, SY11

Student Group	Number of Students	Average Attendance Rate	Average Reading SBA Scaled Score
Caucasian	6,105	96.4%	44.4
Hispanic	15,302	96.1%	38.2
Native American	1,940	94.6%	35.2
Free Lunch	14,455	95.7%	37.2
Reduced-price Lunch	1,352	96.5%	41.2
No FRL	5,732	96.9%	45.2

Source: LFC Analysis

Focused efforts to improve attendance would help overall performance levels. The high-performing schools visited by LFC staff recognized the challenge of improving attendance rates (**Appendix G**). At Griegos Elementary in Albuquerque, for example, the principal described regularly making calls and visiting homes to encourage students to come to school. At Newcomb Elementary in the Central Consolidated School District, bus routes over dirt roads become impassable during inclement weather, but the school used incentives, including food, to increase attendance rates.

From 2003 to 2010, the PED funded a Truancy and Dropout Prevention Program in 13 districts and three charter schools.

Ninety-seven percent of schools received an A for the attendance portion of the A-B-C-D-F school rating system. Schools earned an A for the attendance portion of the school rating system, worth 10 points out of a total of 100 points, for achieving an attendance rate of 95 percent or better. Of 831 preliminary attendance grades, 804 schools earned an A, reducing the incentive for schools to focus on improving attendance rates.

For the SY11 third-grade cohort, mobility within schools in the same district was greater than across districts. Forty-eight percent, or 12,304 students, remained at the same school between kindergarten and third grade and 92 percent, or 23,522 students, remained within the same district.

Recommendations

The Legislature should require districts to annually submit performance-based budgets to the PED, consistent with the General Appropriations Act.

The PED should:
Annually report its process for determining standards-based assessment cut scores and the relative impact on proficiency rates to the Legislature;

Evaluate the impact of bilingual models on the performance of New Mexico's English-language learners;

Require districts to report data on student attendance and identify strategies for improvement as part of its annual performance-based budgeting process; and

Raise attendance criteria in school grading to encourage schools to improve attendance rates.

EARLY EDUCATION IMPROVES PERFORMANCE, BUT LACK OF COORDINATED RESOURCES AND INCONSISTENT QUALITY LIMITS SUCCESS

A high percentage of students show up to kindergarten far behind expectations and are at-risk of academic failure. Children from low-income homes often start out behind their peers and must learn more than a year's worth of academic content each school year to catch up. A brief by the National Center for Children in Poverty confirms these results: preschool-aged children growing up in low-income households score 60 percent below children in the highest socioeconomic group on cognitive tests and know only one-third of the words, 4,000, compared with their middle-income peers, 12,000.

While New Mexico has early learning guidelines, the state lacks a common assessment of kindergarten readiness, making it difficult to compare programs and clearly communicate expectations to parents prior to entering kindergarten. As part of the New Mexico Reads to Lead initiative, the PED plans to spend $3 million for a common formative assessment for kindergarten through third grade.

Some districts, such as Albuquerque Public Schools (APS), have adopted kindergarten readiness assessments, and have found most children are not ready for kindergarten upon entry. APS uses the Kindergarten Developmental Progress Report (KDPR) to assess kindergarten students throughout the school year. The KDPR divides students into four proficiency level, area of need, emergent, nearing proficient, and proficient. In SY08, only 15 percent of students were proficient in language arts at the beginning of kindergarten, compared with a 52 percent proficiency rate as measured with the SBA in the third-grade in SY11. Similar to the SBA, KDPR proficiency rates vary based on the ethnic backgrounds of students entering kindergarten.

Table 8. Fall 2007 Kindergarten Proficiency on the KDPR Assessment

Ethnicity	Area of Need	Emergent	Nearing Proficient	Proficient
Native American	35%	35%	23%	7%
Asian	17%	24%	34%	25%
Black	27%	30%	32%	12%
Hispanic	34%	30%	27%	9%
Caucasian	13%	21%	39%	27%

Source: Albuquerque Public Schools

New Mexico invested $83 million in childcare assistance in 2011 through the Children, Youth, and Families Department, but it is unclear what the program's effects are on children and families. Previous LFC evaluations have noted CYFD does not analyze, and in some cases does not collect, data to determine the effectiveness of the childcare assistance program. A 2009 LFC evaluation of investments in early childhood recommended performance measures could be used to guide data collection, track progress toward increasing the ratings of its providers, and send more children into higher-rated care.

The childcare assistance program subsidizes the cost of childcare for families with incomes at or below 200 percent of the federal poverty level. The program serves approximately 22 thousand children each month, and had a waiting list of nearly 7,000 children in September 2011.

New Mexico, along with the federal government, spends more than $117 million on PreK, Head Start, and special education prekindergarten that help narrow, but not fully close, achievement gaps. For FY13, the state increased FY12 PreK funding levels by 33 percent, appropriating $19.2 million. For FY13, the CYFD and the PED will reimburse sites and districts at a rate of approximately $2,900 per student.

Table 9. Pre-K General Fund Appropriations
(dollars in thousands)

		FY06	FY12	FY13
PED	Appropriation	$2,475.0	$6,292.6	$10,000
	Students	770	2,380	3,103 *
	Sites	24	50	
CYFD	Appropriation	$2,374.5	$8,221.7	$9,200
	Students	770	2,211	2,827*
	Sites	33	70	
Total	Appropriation	$4,849.5	$14,514.3	$19,200
	Students	1,540	4,591	5,930*
	Sites	57	120	

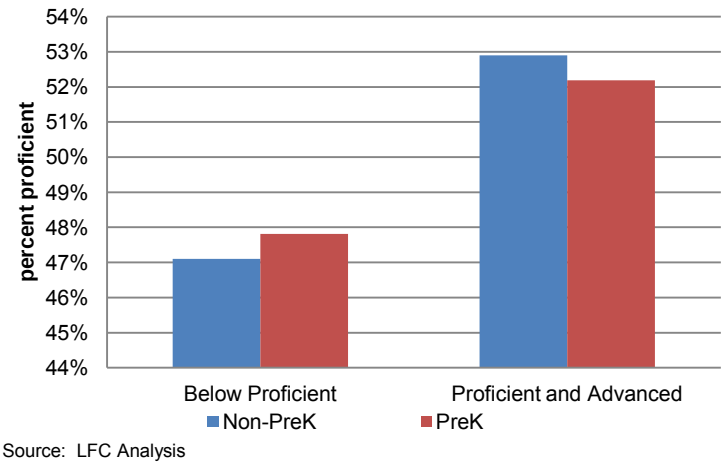
*Estimated Source: PED, CYFD

Many of these non-PreK students, however, received educational programs prior to kindergarten. For example, in FY11, Head Start, a federally funded early childhood program not administered by PED or most school districts, served 5,400 four and five-year olds in New Mexico at a cost of $57 million.

Studies have shown that Head Start is a cost-effective program providing significant short-term impact on the early development of low-income children. A growing body of research in neuroscience, developmental psychology, and economics suggests the earliest years of life are a particularly promising time to intervene in the lives of low-income children. A 2005 federal study of the effectiveness of Head Start programs showed significant positive impacts on behavior, reading, writing, and vocabulary skills. Additionally, a 2008 study by the University of Chicago suggested the federal Head Start program provided significant short-term benefits to its participants that justified the program's costs.

PreK boosts student performance, including third-grade reading scores. In SY11, third-graders who attended state-funded PreK programs were proficient at nearly identical rates, 52 percent, as the overall population of New Mexico third-graders, 53 percent. This is remarkable because these PreK programs served higher percentages of Hispanic, Native American, ELL, and FRL students than the overall population of third-graders. Additionally, the average third-grade reading SBA scaled score for PreK students, 39.6, was higher than for non-PreK students, 39.5. The estimated effect of PreK is even greater when controlling for these demographic differences.

Chart 6. PreK vs. Non-PreK SBA Reading Proficiency Rates, SY11



Source: LFC Analysis

Of the students who attended a PED PreK program, 49 percent were proficient on the third-grade reading SBA and the average scaled score was 38.9. Of students who attended a CYFD PreK program, 56 percent were proficient, averaging a scaled score of 40.4.

Compared with state averages, PED sites enrolled higher percentages of Native American FRL, and ELL students, while CYFD sites enrolled higher percentages of Hispanic students.

Table 10. Demographic Profile of PED and CYFD PreK Students, SY07

	CYFD PreK (778 students)	PED PreK (824 students)	Non-PreK third-Graders, (23,893)
Hispanic	73%	57%	60%
White	17%	9%	25%
Native American	7%	31%	8%
FRL	76%	90%	74%
ELL	15%	32%	20%
Special Education	11%	10%	16%

Source: LFC Analysis

Of the 18,250 third-graders who qualified for free or reduced-price lunch in 2011, 7 percent, or 1,335 students, attended PreK. At FY13 funding levels, it is estimated that 53 percent of four-year olds who qualify for free or reduced-price lunch will participate in PreK or Head Start.

PreK students outscored similar non-PreK students on the reading portion of the SBA when controlling for demographic variables. Compared with a student group made up of similar proportions of Hispanic, Caucasian, Native American, ELL, FRL, and special education students, non-PreK students were estimated to earn an SBA reading scaled score of 39, while PreK students were estimated to earn 40.2 scaled score points (**Appendix H**). This means, on average, participating in PreK is estimated to add 1.2 scaled score points to a student's third-grade reading SBA score. In 2011, 1,654 students, or 6 percent of the 25,495 third-graders with reading SBA scores, were within one point of being considered proficient and 12 percent of all third-graders were within three points of being proficient.

Additionally, the estimated impact of PreK programs on third-grade SBA scores was greatest at PED sites. On the 2011 reading SBA, the estimated effect for PED PreK participants was 1.8 scaled score points; for CYFD PreK participates, the estimated effect was 0.4 scaled score points. These estimates control for ethnicity, ELL, FRL, and special education.

Similar results of the impact of early childhood education programs were found in Albuquerque. A study conducted in 2006 by the Center for Education Research at UNM concluded students who attended an Albuquerque child development center were more likely to graduate from high school, more proficient in reading, and less likely to be classified with a learning disability. The study also determined the initial effects declined during elementary school, highlighting the importance of sustaining intervention services to the neediest students.

As measured by the National Institute for Early Education Research (NIEER) as well as New Mexico's Early Learning Outcomes Assessment (PreK Assessment), New Mexico PreK produces consistent benefits for children. Using standardized early learning assessments, NIEER found that over a three-year period, students showed similar growth from both PED and CYFD PreK programs.

Table 11. NIEER PreK Score Increases, SY06 - SY08

Subject (Test)	Avg. CYFD Point/% Increase	Avg. PED Point/ % Increase
Language (PPVT III Raw Score)	6.27	5.38
Mathematics (WJ III Applied Problems Raw Score)	1.91*	1.44*
Early Literacy (TOPEL Print Knowledge % Correct)	23%*	26%*

*Statistically significant

Source: NIEER

Additionally, for several years, the PED and the CYFD have contracted with the University of New Mexico's Early Childhood Services Center for between $540 thousand and $750 thousand annually to oversee the PreK assessment, maintain a PreK database, mentor PreK teachers, and provide other training and professional development activities.

New Mexico's PreK assessment is one of three state-developed early childhood instruments in the nation and evaluates students on 25 indicators across seven learning domains: physical development, health and well-being; literacy; numeracy; aesthetic creativity; science; self, family, and community; and approaches to learning. Based on the SY11 PreK assessment, 90 percent of PreK students are progressing across the seven measured domains, but it is unclear how many of these students are considered ready for kindergarten.

New Mexico's PreK teachers and administrators value the PreK assessment, but the time dedicated to the assessment is high. A majority of the 94 respondents to an LFC survey agreed the PreK assessment was easy to administer and they received adequate training to administer and interpret the assessment. Seventy-eight percent of respondents felt the assessment informs their teaching, and two-thirds of respondents noted that the information obtained from the assessment helps them communicate student progress to parents.

The statement, "The information I receive from the assessment is worth the time it takes me to administer the assessment," received a relatively low agreement rate, 60 percent, and the lowest scale score of any of the statements, 3.4 out of 5.

PreK practitioners reported spending 25 percent of their school year administering the PreK assessment. The assessment is based on observations of students in class. Survey answers likely took into account those observation hours, which could occur during direct instruction to students. Forty-five percent of PreK teachers spend 100 hours or less administering the assessment, while 16 percent spend 400 hours or more, with an average of 200 hours. Assuming a six-hour school day, this equates to over six weeks of instructional time, or 25 percent of the 800-hour school year.

Table 12. Annual Number of Hours Spent Administering PreK Assessments

Hours Spent Administering PreK Assessment	% of respondents
100 or less	45%
101-200	20%
201-300	8%
301-400	13%
401-500	13%
501+	3%

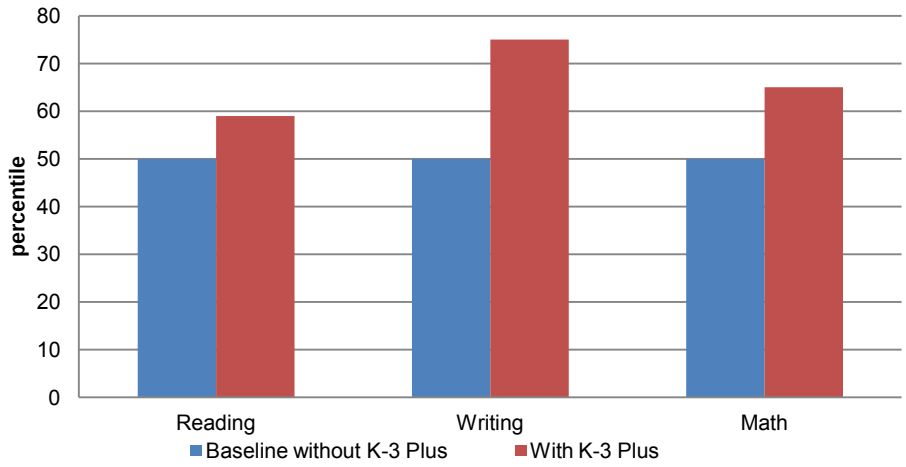
Source: LFC Survey

New Mexico's K-3 Plus program is making positive differences for the neediest students. K-3 Plus was developed in 2007 as a pilot project to narrow the achievement gap by extending the school year by 25 days at schools with more than 85 percent of students qualifying for free or reduced-price lunch. In FY13, New Mexico appropriated $11 million to the K-3 Plus program, more than double the FY12 appropriation of $5.3 million.

In SY08, 2,491 students registered for K-3 Plus from 44 schools across 15 districts. By SY12, the number of registered students doubled to 4,941 students from 50 schools in 15 districts. Districts received $800 per student for 4,564 of these students who met the state's minimum attendance requirement of 18 days. For the 2012 summer session, the Legislature appropriated funding for 9,600 students at $1,100 per student. PED received applications for 9,295 students, 25 percent of the estimated 41 thousand eligible students, from 75 schools in 20 districts and one state charter school.

A 2011 evaluation of New Mexico's K-3 Plus program conducted by Utah State University found positive effects on third-grade reading, writing, and math SBA performance and estimated the benefits from reduced grade retention and remediation services offset all K-3 Plus costs.

Chart 7. Effect of K-3 Plus on SBA Percentile Rank, SY10



Source: Utah State University

Researchers from Utah State University are currently evaluating the effects of K-3 Plus with a control group to minimize potential selection bias; preliminary results are expected later this year.

The state's K-3 Plus program effectively targets high-needs students. In SY11, the demographic profile of the 2,251 third-grade students enrolled in K-3 Plus before the second or third-grade school year were as follows:

- 96 percent qualified for free or reduced-price lunch, compared with 72 percent statewide;
- 79 percent were Hispanic, compared with 60 percent statewide;
- 14 percent were Native American, compared with 9 percent statewide;
- 45 percent were English-language learners, compared with 20 percent statewide; and
- 13 percent had been retained at least once, compared with 10 percent statewide.

Consistent with the risk-factors of this group of K-3 Plus students, performance is lower than state averages: only 33 percent of third-graders who were enrolled in K-3 Plus for one year were proficient or above on the reading SBA, compared with 53 percent of all third-grade students.

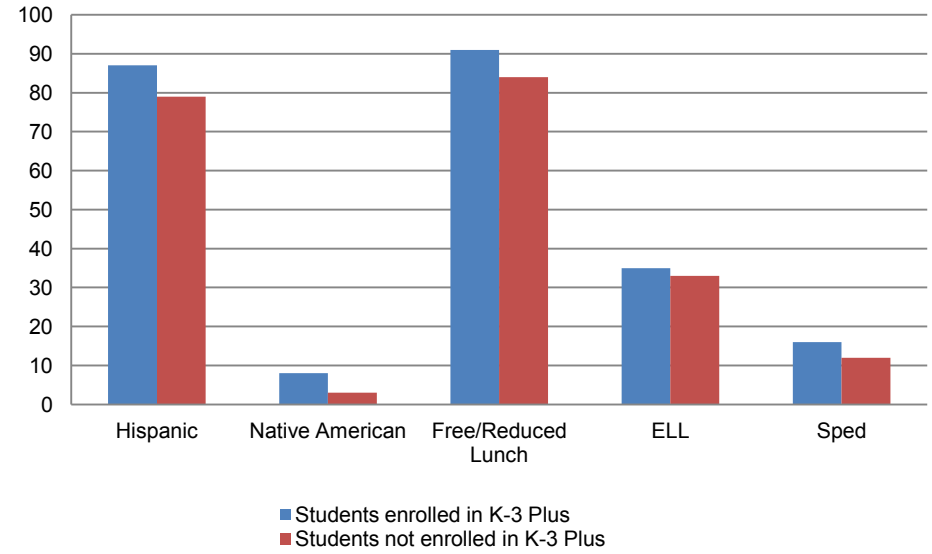
Students who receive two years of K-3 Plus outperform students who attend one year of K-3 Plus. For the 1,603 students who attended one year of K-3 Plus, the average scaled score was 34.7, and for the 665 students who attended two years, the average scaled score increased to 35.1. Controlling for student demographics, students who attended K-3 Plus a second year are estimated to score 0.8 scaled score points higher than students who attended K-3 Plus for one year.

Schools are effectively targeting K-3 Plus to their neediest students. Similar trends exist even within schools that offered K-3 Plus. At seven randomly sampled schools with K-3 Plus, a higher percentage of the students enrolled in the program were Hispanic, Native American, ELL, FRL, or qualified for special education services than for the students at those schools that did not enroll in K-3 Plus.

LCPS' Joint Multi-Age Primary Program (JUMP)

Las Cruces Public Schools is adopting a new approach to ensure at-risk students are able to read by the third grade. The JUMP program allows students in grades K-2 to progress continuously in flexible groups at their own pace by attending an ungraded multi-age classroom. JUMP students will also participate in the district's K-3 Plus program. JUMP targets 14 classrooms in seven elementary schools. The program purposefully places students of different ages together in the same classroom based on ability, while still supporting an individualized, continuous-progress instructional model. LCPS will implement and evaluate the program beginning in SY13.

Chart 8. Student K-3 Plus Demographics at Seven Schools, SY11



Source: LFC analysis

Few at-risk students have access to a full continuum of early education programs, despite the need for extra learning time. In SY12, only 13 schools in New Mexico offered PreK and K-3 Plus. Of 25,495 third-graders with SBA reading scores in 2011, only 81, or 0.03 percent, attended PreK and two years of K-3 Plus. Eighty of these students, or 99 percent, were FRL, 79 percent were Hispanic, 15 percent were Native American, and 41 percent were ELL. Of these 81 students, 49 percent were proficient or advanced on the 2011 reading SBA with an average scaled score of 37.8.

Controlling for ethnicity, ELL, special education, and FRL, students who receive all three programs are estimated to score between 1.5 and 3.8 scaled score points higher than students who do not receive these programs.

Better attention is needed on regular performance reporting and adherence to quality of K-3 Plus program implementation. With the exception of programs that did not intend to offer the required number of days of service, all districts and charter schools that have applied have received K-3 Plus funding. While the application requires plans for dedicated reading and math blocks as well as intervention services for students in the lowest quartile, the PED's ability to ensure proper implementation has been limited.

An evaluation of K-3 Plus in Albuquerque, for example, found programs were more successful if K-3 Plus teachers were paired with the same students they were teaching in the upcoming school year. Similarly, a RAND Corporation study identified maximizing quality, enrollment, and attendance as critical elements to achieving benefits from summer learning programs. In accordance with statute, the PED will disseminate these and other best practices, form a K-3 Plus advisory committee, and promulgate rules for application requirements.

Recommendations

The PED should:
Track student enrollment in PreK, Head Start, or other pre-school programs in the student information database (STARS); also, collect New Mexico PreK assessment data in STARS;

Coordinate allocation of PreK and K-3 Plus resources to increase the number of students who receive the full benefit of both programs; and

Increase oversight and accountability of K-3 Plus with a particular focus on best practices such as addressing student attendance, aligning curriculum, providing remediation and intervention, and matching students with their upcoming teachers.

The PED and the CYFD should:
Consider alternative PreK assessments based on cost effectiveness, time required for administration, alignment with the common core kindergarten standards, and ability for comparisons with other states.

Report student PreK performance as part of the Accountability in Government Act.

EARLY LITERACY INITIATIVES, SUCH AS RETENTION AND READING FIRST, HAVE PRODUCED MIXED RESULTS

To boost performance, many states have implemented mandatory retention policies. Generally, states with such policies retain third-graders reading below proficient and provide remediation and intervention to increase the number of proficient third-graders. In 2002, Florida instituted a statewide third-grade retention policy. Several states have since followed, including Arizona, Oklahoma and, most recently, Indiana. Many other states, including New Mexico, have recently considered similar legislation.

Mandatory retention policies are typically coupled with increased intervention, making it difficult to determine whether retention policies, interventions, or the combination of both are impacting student achievement. Students retained through a state- or district-wide policy typically receive additional support to increase reading performance, including summer reading camps and tutoring during the school year. Also, research typically relies on standardized test data that does not isolate the effects of retention, making it difficult to isolate the causes of changing student achievement.

A study of Chicago Public School's third-grade retention policy, for example, found retained students were unaffected one year after retention and their achievement was 6 percent lower than low-achieving peers who were not retained. In Florida, retained third-graders slightly outperformed socially promoted students in reading in the first year after retention and those gains increased in the second year. Another analysis found retained third-graders caught up with their peers in fourth grade, but those gains had largely been lost by sixth grade.

Retained students are at an increased risk of dropping out of school, earning less income, and engaging in crime. Retained students are two to 11 times more likely to drop out of school. In 2010, the unemployment rate for those without a high school diploma was 5 percent higher than for workers with only a high school diploma. High school graduates earned 41 percent more than those without diplomas. Additionally, high school dropouts in the U.S. are 3.5 times more likely to be incarcerated, and in New Mexico, three-quarters of state prison inmates are high school dropouts.

The Legislature has recently considered revisions to New Mexico's remediation and retention laws. Current law allows parents of students in grades one through eight to refuse to allow their child to be retained for one year. No comprehensive data is available to determine how many retention recommendations are overridden by parents. Schools are then responsible for developing and implementing academic improvement plans, and if the student does not make sufficient progress, that student may be retained for one year.

Recently considered legislation would require retention of third-grade students scoring at beginning steps, the lowest proficiency level on the SBA. Of the 25,495 third-graders with valid reading SBA scores in 2011, 5,628, or 22 percent, met this criteria; 6,392 students, or 25 percent, were nearing proficient, but would not be mandatorily retained. Consistent with similar legislation in other states, students would be exempt from retention for proficient scores on an alternate assessment, completion of a portfolio, English-language learner status, special education status, or previous retention in kindergarten, first, or second grades.

Approximately 8 percent of 2011 third-graders, or 1,923 students, would have been eligible for retention under recently considered legislation. Of the 5,628 students who scored at beginning steps in 2011, 3,705 would have been exempted because they had previously been retained, qualified for special education, or were ELL. Some the remaining 1,923 eligible students would have likely been exempted through alternate assessments or portfolios.

For these 1,923 students:

- The average reading SBA scaled score was 25, compared with a cut score of 32 for nearing proficient and a statewide average of 39.5;
- 57 percent were males;
- 64 percent were Hispanic;
- 15 percent were Native American;
- 86 percent qualified for free or reduced-price lunch;
- 7 percent attended PreK;
- 10 percent attended K-3 Plus before either their second or third-grade school year;
- the average attendance rate was 95.3 percent; and
- 61 percent changed schools between kindergarten and third-grade, higher than the overall average of 52 percent.

Sixty-five percent of these students came from 10 school districts, with 30 percent from the Albuquerque Public Schools.

Table 13. Third-Graders Potentially Eligible for Retention by District, SY11

District	Number of third-graders eligible for retention	% of eligible
Albuquerque	573	30%
Las Cruces	143	7%
Gallup McKinley	102	5%
Rio Rancho	72	4%
Hobbs	67	3%
Central Consolidated	64	3%
Clovis	63	3%
Farmington	62	3%
Roswell	61	3%
Los Lunas	52	3%
Total	**1,259**	**65%**

On average, providing an additional school year costs the state $7,000, the same amount as PreK and four years of K-3 Plus per student. For the 1,923 third-graders eligible for retention in SY11, this would have cost the state $13.5 million, compounding with each cohort subject to a mandatory third-grade retention policy.

In New Mexico, almost 10 percent of the third-grade class in SY11 was retained between kindergarten and third-grade. The greatest number, 936, were retained in kindergarten, decreasing steadily to 232 retained in the third-grade. The average reading SBA scaled score for all non-retained third-graders in 2011 was 40.2, compared with 32.9 for retained students. Of those retained students, 725, 30 percent, qualified for special education services by the time they were third-graders; 208 of those special education students required high levels of services. The average attendance rate for non-retained students, 96.1 percent, was higher than the average attendance rate for retained students, 95.3 percent.

As is the case nationally, some groups of retained students were over-represented: Hispanics, males, FRL, and ELL. Also, average scaled scores and proficiency rates were lower for retained students than for non-retained students.

Table 14. Retention and Third-Grade SBA Reading Scores, SY11

	Number of Students	Average Scaled Score (Proficient = 40)	% Proficient
Non-retained	23,076	40.2	53%
Retained in Kindergarten	936	33.4	32%
Retained in First	732	31.7	24%
Retained in Second	528	32.5	25%
Retained in Third	232	35.2	34%

Source: LFC analysis

Of the retained third-graders who scored at beginning steps in SY10, only 12 percent were proficient by the end of their second year of third grade. Of the 100 third-graders who scored at beginning steps and were retained in SY10, 25 stayed at beginning steps in SY11, 63 moved up one level to nearing proficient, and 12 moved to proficient. Although many students increased their reading scores in their second year of third grade, overall, only 29 percent moved up to become proficient in reading after repeating the third-grade. Forty-three percent of students retained in third grade in SY10 did not improve a proficiency level or regressed a proficiency level.

Table 15. Retained Third-Graders Moving to Proficient or Above, SY10 and SY11

SY10 Proficiency Level	Number of students retained in SY10	Number of Retained Students Scoring Proficient in SY11	Number of Retained Students Scoring Advanced in SY11	Proficient or Advanced after 2nd year of 3rd grade, SY11
Beginning Steps	100	12	0	12%
Nearing Proficient	89	42	0	47%

Source: LFC Analysis

Students retained in third grade in SY10 increased their SY11 SBA reading scaled score by an average of 8.8 points, but the change in student scores ranged from a decrease of 20 points to an increase of 32. Fourteen percent of students retained in third grade in SY10 saw their SBA reading scores decline or stay the same in SY11. Another 18 percent of students increased their third-grade SBA reading score by four points or less, while 23 percent improved their scores by 15 points or more.

Table 16. Retained Third-Grade SBA Reading Score Change from SY10 to SY11

SBA Reading Point Change from SY10 to SY11	% of Retained Students
≤ 0	14%
1-4	18%
5-8	18%
9-11	15%
12-14	11%
15-18	11%
>18	12%

Source: LFC Analysis

Retained third-graders with the lowest SY10 reading proficiency levels grew at the greatest rates. Third-grade students who scored at beginning steps on the SY10 SBA averaged a 13-point increase on their SY11 SBA reading score and only 5 percent of those students' scores declined. In contrast, third-graders who scored nearing proficient and proficient prior to being retained averaged smaller increases and more frequently had scores that declined in their second year of third grade.

Table 17. Changes in SBA Scores for Retained Third-Graders

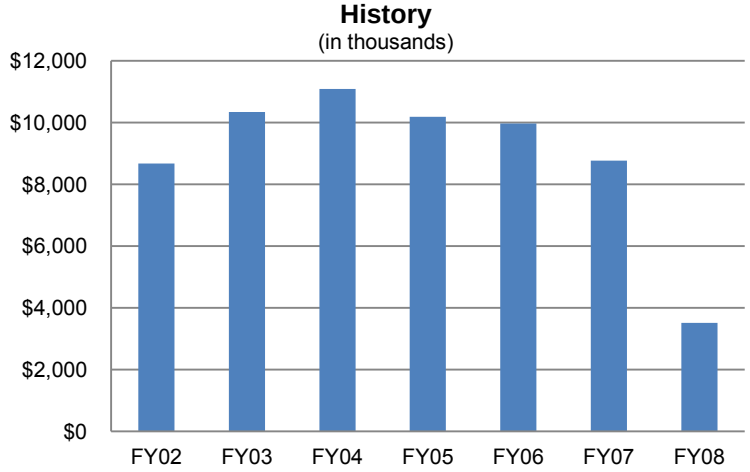
2010 Third-Grade Reading Proficiency Level	Average Score Increase from SY10 to SY11	Percent of Students with Declining Scores	# of students
Beginning Steps	13	5%	110
Nearing Proficiency	6	16%	89
Proficient	4	25%	20

Source: LFC Analysis

Nine percent of the 219 third-graders retained in 2010 were proficient in reading. Twenty students who had scored proficient on the reading SBA were retained as third-graders in 2010. While these students increased their reading scaled scores by an average of 3.5 points in their second year of third grade, 10 percent, or two students, dropped to nearing proficient.

During Reading First, a federally funded program similar to the currently proposed Reads to Lead initiative, third grade reading proficiency rates did not improve. From 2002 to 2008, as part of No Child Left Behind, Congress appropriated $1 billion annually in six-year grants to states to improve early literacy rates. New Mexico received $62.5 million in federal Reading First funds. In FY07, these funds served 19 thousand students in 100 schools across 39 districts in New Mexico.

Chart 9. New Mexico Reading First Funding History

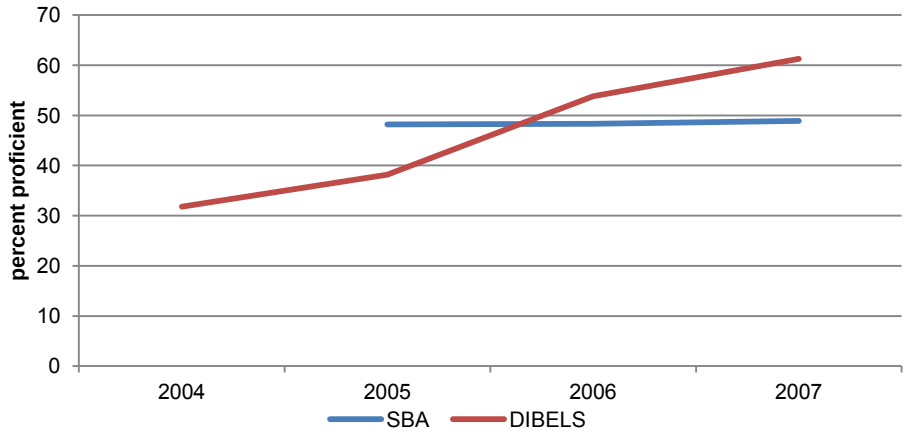


Source: Reading First State Profile

Reading First required states and participating school districts to adopt scientifically based reading programs, provide professional development, use reading coaches, and track students' reading progress using valid and reliable assessments. Congress did not reauthorize the program in 2009, based partially on the U.S. Inspector General's findings of mismanagement and conflicts-of-interest.

While DIBELS scores rose during Reading First, state SBA scores remained flat. As measured by the dynamic indicators of basic early learning skills (DIBELS), New Mexico's fluent third-grader readers increased from 23 percent in 2004 to 61 percent in 2007. Over this same time, however, SBA reading proficiency rates remained flat at 48 percent, highlighting differences between the DIBELS, which measures fluency, and the SBA, which measures comprehension.

Chart 10. Third-Grade SBA Reading and Third-Grade DIBELS Proficiency Data



Source: Reading First Annual Performance Report

In contrast, 11 states increased proficiency rates on their standardized tests while implementing Reading First. Evaluations conducted in 2007 found characteristics of Reading First states that increased both DIBELS and state-based proficiency assessment scores included stable school leadership and principals and teachers with a strong commitment to the Reading First principles.

Table 18. 3rd Grade State Assessment Trends of Schools Participating in Reading First

State Test Scores Increased	Test Scores Remained Flat	State Test Scores Decreased
Arizona	Georgia	Delaware
Connecticut	North Dakota	Florida
Hawaii	Ohio	Massachusetts
Illinois	**New Mexico**	Mississippi
Indiana	Utah	Oregon
Mississippi		
New Jersey		
Pennsylvania		
South Carolina		
West Virginia		
Wyoming		

Source: US Department of Education

Recommendation

The PED should annually report the number of students retained in each grade by district.

STATE, DISTRICT, AND SCHOOL-LEVEL MANAGEMENT PRACTICES CAN HELP SCHOOLS TO MARGINALLY BEAT THE ODDS

Quality teaching is the most important school-based influence on student achievement. Numerous studies have led researchers to conclude, "The most important factor affecting student learning is the teacher" (Wright, Horn, and Sanders, 1997, p. 63). Additionally, poor and minority students show the greatest academic gains when paired with an effective teacher, although nationally these higher-risk students tend to have less experienced teachers with lower licensure levels.

In 2009, the National Council on Teacher Quality (NCTQ) evaluated eight New Mexico teacher preparation programs' admissions standards, reading programs, elementary math programs, and exit standards. Overall, NCTQ found low admissions standards; lack of focus on the science of teaching reading, including poor reading textbook quality; inadequate math preparation, including poor math textbook quality; and a weak exit assessment.

This NCTQ study highlights the importance of both teacher preparation and ongoing professional development. Upcoming LFC evaluations will consider the effect of the state's teacher preparation programs as well as teacher evaluation on student performance as measured by standardized test scores.

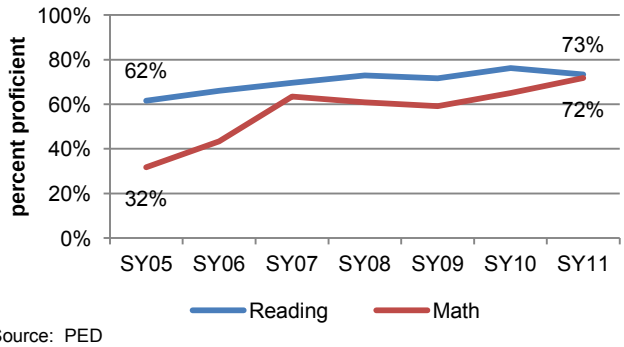
School leadership also has an indirect effect on student learning. Numerous researchers have established that strong leadership has a statistically significant influence on student achievement, accounting for up to 25 percent of the total school effect (Marzano, 2000). One effort to synthesize the descriptions of best practices resulted in the Interstate School Leaders Licensure Consortium (ISLLC) standards, including vision-setting, developing a culture focused on learning, strong management, and collaborating within and outside of the school.

Locally, Gus Benakis, the former principal at Harrison Schmitt Elementary in Silver City, developed his own "Characteristics of Success" consistent with the ISLLC standards:

1. Raise the expectations, clarify the focus;
2. Communicate (Listen!);
3. Be visible (especially the principal);
4. Collaborate K through 5;
5. Identify/ assess student needs early;
6. Retain early (K, 1);
7. Align curriculum (especially in weak areas);
8. Provide professional development;
9. Embrace challenges and acknowledge success; and
10. Take away the excuses and provide necessary resources.

As a result, student proficiency rates have steadily increased over the last six years.

Chart 11. Harrison Schmitt Elementary SBA Proficiency Rates, All Students



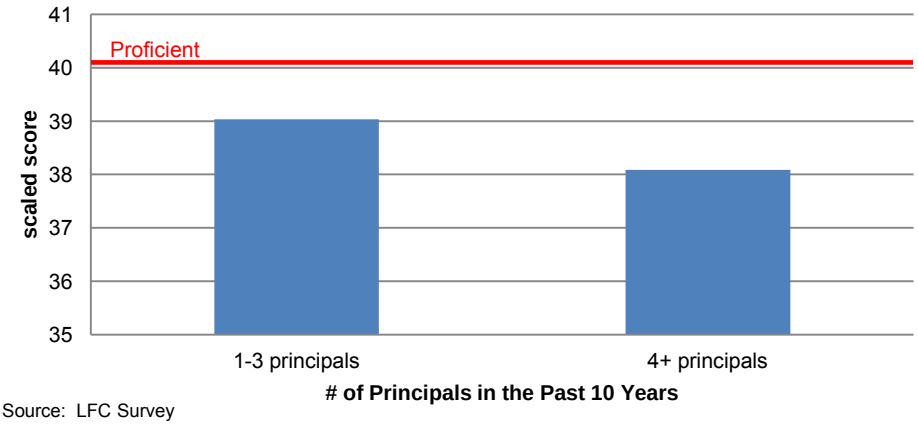
Source: PED

Based on an LFC survey of New Mexico elementary school principals, schools with more principal turnover generally had lower reading scores on the third-grade SBA. Of 89 elementary schools that responded with complete information, those with two or fewer principals over the past 10 years averaged a third-grade SBA reading scaled score of 39, while those with four or more principals averaged a score of 38.

The Importance of Stable Leadership

Of the six high-performing schools LFC staff visited—Dolores Gonzalez, Washington Avenue, Griegos, Mesilla Park, Jaramillo, and Newcomb Elementary—each had principals with tenures of 10 years or more. In contrast, of the six low-performing schools visited, each had at least three and as many as five principals in the last 10 years.

Chart 12. Number of Principals by School in the Last 10 Years and Average Third-Grade Reading Scaled SBA Score



Source: LFC Survey

Of those 89 respondents, 29 percent of schools had four or more principals in the past 10 years.

Table 19. Principal Turnover

Number of principals by school in the last 10 years	% of Respondents
1	26%
2	21%
3	24%
4	13%
5 or more	16%

Source: LFC Survey Data

Elements within the current administrator licensure requirements and minimum salary structure act as obstacles to the supply of qualified school leaders in New Mexico. Over the last five years, the number of school leaders prepared by New Mexico's five higher education institutions has decreased by 38 percent, from 138 to 86.

One barrier to administration identified by numerous district leaders is the length of time required to earn an administrative license. To progress to a level III administrative license in New Mexico requires a minimum of six years teaching experience or seven years for out-of-state applicants. In contrast, neighboring states require fewer years. For example, Texas and Oklahoma require only two years and Colorado and Arizona each require three years. Additionally, Colorado and Oklahoma offer alternative licensure for promising leaders without teaching experience.

A second barrier cited is that per contract day, minimum elementary principal salaries are less than minimum level III teaching salaries. State statute established a minimum annual salary of $50 thousand for level III teachers and $60 thousand for elementary principals. Based on typical contract lengths for each position, level III teachers earn a minimum of $278 per day, while elementary principals earn $273 per day.

High-quality implementation of best practices impacts student growth as measured by the SBA. For example, 89 of the 167 respondents to an LFC survey, 53 percent, indicated having reading coaches. How those coaches spend their time, however, directly relates to student achievement. Similarly, effective use of data to drive instructional decisions is affected by fidelity of implementation. Like the six high-performing schools visited for this evaluation, Harrison Schmitt Elementary is a school that meaningfully uses data (see **Appendix I**).

Use of Measures of Academic Progress (MAP) as a short-cycle assessment in kindergarten and third grade related to higher student performance. Based on results from an LFC survey, in kindergarten and third grade, use of MAP, a national computer-adaptive assessment produced by the Northwest Evaluation Association, has a statistically significant positive correlation with schools' average scaled scores. Of 167 responding schools to the LFC survey, 20 percent used MAP in kindergarten and 38 percent used MAP for third-graders.

Table 20. Use of MAP by Grade Level

Grade	Percent Using MAP
Kindergarten	20%
1st grade	23%
2nd Grade	36%
3rd Grade	38%

Source: LFC Survey Data

Reading coaches spending time analyzing data is related to school performance. While other functions performed by reading coaches did not correlate with school performance, reading coaches performing data analysis was found to have a positive relationship. Of 89 schools surveyed with reading coaches, 29 percent indicated that their reading coach performed some data analysis. Those reading coaches spent between 5 percent and 45 percent of their time on data analysis.

The use of DIBELS on a weekly or monthly basis had a positive correlation with higher SBA scores. While the use of DIBELS alone was not connected to school success, schools that conducted the DIBELS assessment on a weekly basis had SBA scores higher than those that did not conduct the assessment at this frequency. Three percent of schools surveyed conducted DIBELS assessments weekly.

Conducting DIBELS assessments monthly was also connected with higher school performance. Eight percent of schools surveyed performed DIBELS on a monthly basis. The use of DIBELS on a bi-weekly, quarterly, or semi-annual basis was not correlated with higher SBA scores.

Jaramillo's Fidelity to Data

Jaramillo Community School in Belen uses DIBELS to regularly monitor student progress and move students into appropriate reading groups based on reading levels. Everyone from the principal to the literacy coach to teachers participates in this decision-making process and can readily describe what each student needs to progress as a reader.

DIBELS scores correlate with performance on the SBA. Statistically significant correlations were found between student scores on DIBELS oral reading fluency (ORF) and nonsense word fluency (NWF) assessments and the SBA.

As described earlier, however, during Reading First, improvements in DIBELS scores did not result in increased SBA proficiency rates. One possible explanation is that based on a sample of 3,400 New Mexico third-graders, students have to score significantly above the DIBELS benchmark to be considered proficient on the third-grade reading SBA. For example, kindergarteners are considered low-risk, the highest level on the DIBELS, with a score of 25. Students who went on to be proficient on the SBA, however, averaged a 78 on the kindergarten DIBELS.

Table 21. DIBELS Scores of SBA Proficient Students Compared with Low-Risk Benchmarks, SY11

Grade and DIBELS Assessment	Low-Risk Benchmark	Avg. Score for SBA Proficient and Above	Avg. Overall Score
Kindergarten (NWF)	25	78	69
1st Grade (ORF)	40	65	46
2nd Grade (ORF)	90	107	79
3rd Grade (ORF)	110	112	95

Source: PED

Recommendations

The Legislature should revise statute to require a minimum of three years of teaching, or level II licensure, to obtain an administrative license.

The PED should:
Require districts to report data on principal and teacher turnover and identify strategies for improvement as part of an annual performance-based budgeting process.

Adopt statewide, short-cycle assessments in grades K-3 that align to the common core standards, measure growth of all students at least three times per year, can be used more frequently to monitor the progress of higher-needs students, and allow comparisons with other states.

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

HANNA SKANDERA
SECRETARY-DESIGNATE OF EDUCATION

SUSANA MARTINEZ
Governor

Public Education Department

Formal Response to the
Report to the Legislative Finance Committee:
Early Literacy
New Mexico Public Education Department

June 13, 2012

The Public Education Department (PED) wishes to thank the Legislative Finance Committee (LFC) for the opportunity to review and provide feedback on the draft *Report to the Legislative Finance Committee* addressing the topic of early literacy, including the Prekindergarten, K-3 Plus, and third grade programs. Informal feedback was provided to LFC through meetings and email communications taking place May 22 – June 1, 2012. The purpose of this document is to provide a response, as well as a summary, by PED.

General Observations

- It is recommended to exercise caution when using the terms "influence" and "impact" as evidence is not always available to assert causation.
- Consideration of what criteria were used to answer the questions as well as statistical and substantive evidence to describe meaningful differences.
- Inconsistent use of data: Different data were used to make different conclusions for different grade levels.
- It is recommended to ensure that students with disabilities are represented throughout the data presented in the report.

Feedback from PED Program Staff

The following feedback items were previously shared with LFC:

- With regard to third grade, PED recommended that LFC provide clarification for data graphics, including the addition of a color code key for a socioeconomic status map and enhance the chart on the third grade achievement gap for which LFC provided revisions based on feedback.
- The draft report indicated that "New Mexico lacks statewide kindergarten readiness standards and a common assessment of readiness…" PED shared that there are Early Learning Guidelines for prekindergarten programs providing the skills children need to have before entering kindergarten. All PreK programs are required to use the Early Learning Guidelines. LFC revised the draft report to include this language. Further, Early Learning Guidelines have been extended through kindergarten.
- The draft report included language that "PED and CYFD should consider alternative PreK assessments…in alignment with the Common Core kindergarten standards." The current PreK assessment was analyzed to crosswalk with the Common Core State Standards. The crosswalk was provided to LFC on June 2, 2012 per its staff request.
- With regard to K-3 Plus, the draft report indicated that "No sites, however, offered the program to all four grade levels." As follow-up, PED provided the data to LFC staff demonstrating that the majority of sites offer the K-3 plus program to all four grade levels (kindergarten, first, second, and third grades).
- The draft stated that "PED's assignment of school grades favored schools with fewer poor students". PED indicated to LFC staff that the correlation between free and reduced priced lunch status under AYP was -.57, versus -.417 for status under school grades, and .090 for growth under school grades. The school grading system has taken positive steps in accurately holding schools accountable.

Once again, we appreciate the opportunity to respond to the draft report. PED is committed to continuous quality improvement and constantly striving to improve the outcomes of our students and programs. We remain accountable to those we serve and to all stakeholders involved in public education in our state. The report that the LFC has provided is seen as a resource and a tool to build our capacity and continue to improve through effective use of data, evaluation, best practices and innovative strategies.

APPENDIX A: PROJECT INFORMATION

Evaluation Objectives.

1. Performance. Determine reading proficiency rates over time and relationships to student demographics.
2. Finance. Evaluate spending patterns, programs, and practices the state and districts use to finance early literacy.
3. Policy and Programming. Analyze best practices for accelerating student achievement in literacy.

Evaluation Procedures.

- Reviewed best practices in early literacy, paying particular attention to statewide student retention measures and unique financing incentives.
- Reviewed state, district, and school-level student performance data and student demographic data.
- Selected six over-performing and six under-performing elementary schools based on the difference between expected and actual SBA scores given the schools' percentage of students qualifying for free or reduced-price lunch, geographic location, and school size:
  - Chaparral Elementary School (Santa Fe)
  - Conlee Elementary School (Las Cruces)
  - Dolores Gonzales Elementary School (Albuquerque)
  - Emerson Elementary School (Albuquerque)
  - Griegos Elementary School (Albuquerque)
  - Henry T. Jaramillo Community School (Belen)
  - Hernandez Elementary School (Espanola)
  - Mesilla Park Elementary School (Las Cruces)
  - Newcomb Elementary School (Central Consolidated)
  - Ojo Amarillo Elementary School (Central Consolidated)
  - Sunset Elementary School (Roswell)
  - Washington Avenue Elementary School (Roswell)

  On site visits to these twelve elementary schools, evaluators observed over 100 classrooms, interviewed 75 teachers, principals, reading coaches, and district-level leadership.
- Analyzed cohort data for 25,495 third-graders in SY11 with valid standards-based assessment reading scaled scores. Student-level data included demographic information, enrollment in PreK, enrollment in K-3 Plus, and attendance data in grades one through three.
- Electronically surveyed elementary principals statewide regarding best practices in reading.
- Electronically surveyed CYFD and PED PreK practitioners with a particular focus on the PreK assessment.
- Reviewed and determined the cost-effectiveness of current early literacy expenditures.
- Reviewed applicable laws and regulations; LFC file documents, including all available project documents; relevant performance reviews from other states; and performance measures.

Evaluation Team.

Michael Weinberg, Lead Program Evaluator
Matthew Pahl, Program Evaluator

Authority for Evaluation. LFC is authorized under the provisions of 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. LFC is also authorized to make recommendations for change to the Legislature. In furtherance of its statutory responsibility, LFC may conduct inquiries into specific transactions affecting the operating policies and cost of governmental units and their compliance with state laws.

Exit Conferences. The contents of this report were discussed with PED on May 22, 2012.

Report Distribution. This report is intended for the information of the Office of the Governor; the Public Education Department; the Children, Youth, and Families 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.

Charles Sallee
Deputy Director for Program Evaluation

APPENDIX B: PUBLIC EDUCATION DEPARTMENT PERFORMANCE REPORT CARD, THIRD QUARTER, FY12

Performance Overview: In general, little or no consistent public school performance data is available during the year. Performance measures for public school support provide an annual snapshot of student performance at the end of each school year. Student performance at the end of FY11 generally failed to show improvement over FY10. Data from the FY11 administration of the New Mexico Standards Based Assessment shows decreases in statewide proficiency over FY10: a decrease of 3.4 percentage points in reading, 0.4 percentage points in math, and 4.2 percentage points in science. Based on FY11 assessment data, 50.2 percent of students scored below proficient in reading, 58.2 percent of students scored below proficient in math, and 58 percent of student scored below proficient in science.

The Public Education Department reports an increasing number of schools failing to make adequate yearly progress (AYP). Based on assessment results from the 2011 school year, 720, or 86.6 percent of all schools failed to make AYP and are in the school improvement cycle for the 2011 school year. This is an increase of 86 schools over the 2010 school year. Since 2005, the number of schools failing to make AYP has increased 73.1 percent. It is important to note that student achievement is a better indicator of academic success.

The department notes a 4.3 percent decrease in FY11's four-year cohort graduation rate, from 67.3 percent to 63 percent, for freshmen entering high school in 2007 and graduating in 2011. Graduation rate reporting methodologies delay graduation rate reporting by more than a year. However, a high note in student performance, the percentage of recent high school graduates requiring remedial courses in institutions of higher education showed positive progress, dropping from 47.1 percent in FY10 to 46.2 percent in FY11. Student achievement continues to indicate the need for programs that engage students, target struggling students, keep students in school, and better prepare students for college or the workplace.

	Budget: $2,338,422.0 FTE:	FY10 Actual	FY11 Actual	FY12 Target	Q1	Q2	Q3	Q4	Rating
1	Percent of fourth-grade students who achieve proficiency or above on standard-based assessments in reading*	51.4%	46.5%	78%	Reported Annually	Reported Annually	Reported Annually		-
2	Percent of eighth-grade students who achieve proficiency or above on the standards-based assessments in reading*	60.5%	53.3%	76%	Reported Annually	Reported Annually	Reported Annually		-
3	Percent of fourth-grade students who achieve proficiency or above on the standards-based assessments in mathematics*	45.4%	44.4%	77%	Reported Annually	Reported Annually	Reported Annually		-
4	Percent of eighth-grade students who achieve proficiency or above on the standard-based assessments in mathematics*	39.2%	40.8%	74%	Reported Annually	Reported Annually	Reported Annually		-
5	Percent of recent New Mexico high school graduates who take remedial courses in higher education at two-year and four-year schools*	47.1%	46.2%	40%	Reported Annually	Reported Annually	Reported Annually		-
6	Current year's cohort graduation rate using four-year cumulative method*	67.3%	63%	75%	Reported Annually	Reported Annually	Reported Annually		-
7	Annual percent of core academic subjects taught by highly qualified teachers, kindergarten through twelfth grade	99.5%	97.1%	100%	Reported Q2	98.4%	99.3%		G
Program Rating		Y	R						Y

Comments: For FY13, the Legislature appropriated $2.5 million to the PED for short-cycle assessments to be administered in fourth through tenth grades. Short-cycle assessments are designed to assist in making instructional decisions and can be used to indicate student growth within a school year. To be meaningful, implementation should consider mandatory reporting to the Public Education Department (PED) at least three times a year, allowing policy makers access to data more

than once annually.

Improving Student Achievement and Closing the Achievement Gap: Student achievement at the end of the 2010-2011 school year failed to achieve significant gains. Proficiency targets have generally been set unrealistically high, historically. Proficiency targets for FY13 have been adjusted downward to reflect reasonable student achievement growth over time. Despite generally failed student achievement targets by all subgroups, the achievement gap continues to persist in New Mexico, and continues to widen for economically disadvantaged students and English-language learners. The department does not currently report any performance measures for any student subgroups. To better assess the achievement gap, the PED should consider reporting proficiency results by race/ethnicity and additionally report results for economically disadvantaged students and English-language learners.

Teacher Quality: Despite having a "highly qualified" teacher workforce, improvement in student achievement continues to progress slowly. The PED has agreed to reform the state's teacher evaluation system to measure the effect teachers have on student learning as measured by academic growth in exchange for the federal government granting New Mexico a flexibility waiver from requirements No Child Left Behind. While the Legislature failed to reach consensus on teacher evaluation legislation during the 2011 and 2012 sessions, the department will seek to establish the new system in regulations. The Legislature made a $1 million special appropriation to PED to implement a teacher evaluation system based on student achievement growth.

Suggested Performance Measure Improvement

Performance measures for AYP reporting should be phased out in FY14 in exchange for measures aligned with the state accountability A through F rating system, consistent with the federal No Child Left Behind Waiver. Additionally, measures related to teacher and school leader effectiveness ratings should be included for FY14 as more information becomes available from the department regarding the new state evaluation system that will be implemented as part of the federal waiver. Performance measures should be added for student subgroups.

APPENDIX C: SBA READING PERCENT PROFICIENT AND ABOVE

District	School	SY11	SY12	District	School	SY11	SY12
Alamogordo	Buena Vista Elementary	75%	57%	Albuquerque	Coronado Elementary	62%	43%
Alamogordo	Heights Elementary	41%	42%	Albuquerque	Corrales Elementary	78%	71%
Alamogordo	High Rolls Mountain Elementary			Albuquerque	Corrales International Charter	76%	71%
Alamogordo	Holloman Intermediate	75%	72%	Albuquerque	Dennis Chavez Elementary	83%	76%
Alamogordo	La Luz Elementary	75%	60%	Albuquerque	Dolores Gonzales Elementary	51%	45%
Alamogordo	North Elementary	45%	44%	Albuquerque	Double Eagle Elementary	82%	92%
Alamogordo	Oregon Elementary	50%	46%	Albuquerque	Douglas Macarthur Elementary	60%	75%
Alamogordo	Sacramento Elementary	47%	22%	Albuquerque	Duranes Elementary	35%	44%
Alamogordo	Sierra Elementary	72%	82%	Albuquerque	East San Jose Elementary	28%	39%
Alamogordo	Yucca Elementary	64%	67%	Albuquerque	Edmund G Ross Elementary	49%	45%
Albuquerque	A Montoya Elementary	57%	45%	Albuquerque	Edward Gonzales Elementary	37%	40%
Albuquerque	Acoma Elementary	41%	52%	Albuquerque	El Camino Real Academy Charter	16%	43%
Albuquerque	Adobe Acres Elementary	31%	34%	Albuquerque	Emerson Elementary	4%	21%
Albuquerque	Alameda Elementary	49%	66%	Albuquerque	Eubank Elementary	20%	34%
Albuquerque	Alamosa Elementary	48%	44%	Albuquerque	Eugene Field Elementary	46%	41%
Albuquerque	Alice King Community Charter	75%	84%	Albuquerque	Family School	90%	87%
Albuquerque	Alvarado Elementary	60%	49%	Albuquerque	Georgia O Keeffe Elementary	82%	84%
Albuquerque	Apache Elementary	51%	55%	Albuquerque	Governor Bent Elementary	41%	40%
Albuquerque	Armijo Elementary	25%	31%	Albuquerque	Griegos Elementary	78%	59%
Albuquerque	Arroyo Del Oso Elementary	62%	78%	Albuquerque	Hawthorne Elementary	33%	27%
Albuquerque	Atrisco Elementary	57%	50%	Albuquerque	Hodgin Elementary	24%	45%
Albuquerque	Bandelier Elementary	68%	77%	Albuquerque	Hubert H Humphrey Elementary	86%	68%
Albuquerque	Barcelona Elementary	61%	46%	Albuquerque	Inez Elementary	54%	61%
Albuquerque	Bel Air Elementary	33%	39%	Albuquerque	John Baker Elementary	81%	70%
Albuquerque	Bellehaven Elementary	38%	63%	Albuquerque	Kirtland Elementary	62%	40%
Albuquerque	Carlos Rey Elementary	37%	31%	Albuquerque	Kit Carson Elementary	38%	30%
Albuquerque	Chamiza Elementary	58%	68%	Albuquerque	La Luz Elementary	41%	33%
Albuquerque	Chaparral Elementary	63%	61%	Albuquerque	La Mesa Elementary	54%	41%
Albuquerque	Chelwood Elementary	48%	51%	Albuquerque	Lavaland Elementary	32%	38%
Albuquerque	Christine Duncan Heritage Academy Charter	9%		Albuquerque	Lew Wallace Elementary	65%	42%
Albuquerque	Cochiti Elementary	53%	54%	Albuquerque	Longfellow Elementary	49%	38%
Albuquerque	Collet Park Elementary	59%	72%	Albuquerque	Los Padillas Elementary	33%	44%
Albuquerque	Comanche Elementary	64%	59%	Albuquerque	Los Ranchos Elementary	34%	33%

District	School	SY11	SY12	District	School	SY11	SY12
Albuquerque	Lowell Elementary	31%	22%	Albuquerque	Wherry Elementary	32%	24%
Albuquerque	Manzano Mesa Elementary	50%	64%	Albuquerque	Whittier Elementary	33%	41%
Albuquerque	Marie M Hughes Elementary	68%	52%	Albuquerque	Zia Elementary	46%	64%
Albuquerque	Mark Twain Elementary	30%	58%	Albuquerque	Zuni Elementary	71%	65%
Albuquerque	Maryann Binford Elementary	34%	42%	Animas	Animas Elementary	60%	60%
Albuquerque	Matheson Park Elementary	51%	50%	Artesia	Central Elementary	73%	69%
Albuquerque	McCollum Elementary	49%	62%	Artesia	Hermosa Elementary	67%	40%
Albuquerque	Mission Avenue Elementary	49%	46%	Artesia	Penasco Elementary		
Albuquerque	Mitchell Elementary	54%	54%	Artesia	Roselawn Elementary	42%	57%
Albuquerque	Monte Vista Elementary	68%	71%	Artesia	Yeso Elementary	58%	47%
Albuquerque	Montessori Of The Rio Grande Charter	74%	71%	Artesia	Yucca Elementary	69%	55%
Albuquerque	Montezuma Elementary	37%	38%	Aztec	Lydia Rippey Elementary	61%	50%
Albuquerque	Mountain Mahogany Charter	77%	81%	Aztec	McCoy Avenue Elementary	49%	57%
Albuquerque	Mountain View Elementary	43%	34%	Aztec	Mosaic Academy Charter	45%	38%
Albuquerque	Navajo Elementary	29%	29%	Belen	Dennis Chavez Elementary	70%	60%
Albuquerque	North Star Elementary	87%	90%	Belen	Family School	90%	80%
Albuquerque	Onate Elementary	64%	59%	Belen	Gil Sanchez Elementary	65%	59%
Albuquerque	Osuna Elementary	75%	69%	Belen	Jaramillo Elementary	55%	61%
Albuquerque	Painted Sky Elementary	49%	58%	Belen	La Merced Elementary	47%	55%
Albuquerque	Pajarito Elementary	35%	31%	Belen	La Promesa Elementary	41%	43%
Albuquerque	Petroglyph Elementary	77%	71%	Belen	Rio Grande Elementary	45%	51%
Albuquerque	Reginald Chavez Elementary	50%	47%	Bernalillo	Algodones Elementary	48%	48%
Albuquerque	Rudolfo Anaya Elementary	43%	41%	Bernalillo	Cochiti Elementary	30%	16%
Albuquerque	S Y Jackson Elementary	83%	78%	Bernalillo	Placitas Elementary	59%	77%
Albuquerque	San Antonito Elementary	88%	78%	Bernalillo	Santo Domingo Elementary	20%	26%
Albuquerque	Sandia Base Elementary	65%	66%	Bernalillo	Willanna D Carroll Elementary	43%	40%
Albuquerque	Seven Bar Elementary	69%	57%	Bloomfield	Blanco Elementary	33%	52%
Albuquerque	Sierra Vista Elementary	66%	65%	Bloomfield	Central Primary	45%	47%
Albuquerque	Sombra Del Monte Elementary	60%	52%	Capitan	Capitan Elementary	75%	56%
Albuquerque	Sunset View Elementary	68%	71%	Carlsbad	Craft Elementary	65%	59%
Albuquerque	Susie R Marmon Elementary	45%	40%	Carlsbad	Hillcrest Elementary	45%	45%
Albuquerque	Tierra Antigua Elementary	63%	69%	Carlsbad	Jefferson Montessori Charter	68%	52%
Albuquerque	Tomasita Elementary	41%	26%	Carlsbad	Joe Stanley Smith Elementary	60%	58%
Albuquerque	Valle Vista Elementary	24%	28%	Carlsbad	Monterrey Elementary	64%	73%
Albuquerque	Ventana Ranch Elementary	69%	58%	Carlsbad	Pate Elementary	47%	46%

District	School	SY11	SY12	District	School	SY11	SY12
Carlsbad	Puckett Elementary	62%	63%	Cuba	Cuba Elementary	38%	46%
Carlsbad	Riverside Elementary	82%	94%	Deming	Bataan Elementary	40%	49%
Carlsbad	Sunset Elementary	66%	62%	Deming	Bell Elementary	27%	32%
Carrizozo	Carrizozo Elementary	45%	55%	Deming	Chaparral Elementary	46%	56%
Central Consolidated	Eva B Stokely Elementary	31%	50%	Deming	Columbus Elementary	47%	44%
Central Consolidated	Kirtland Elementary	55%	49%	Deming	Memorial Elementary	35%	45%
Central Consolidated	Mesa Elementary	26%	45%	Deming	Ruben S Torres Elementary	29%	27%
Central Consolidated	Naschitti Elementary	59%	30%	Des Moines	Des Moines Elementary		
Central Consolidated	Newcomb Elementary	61%	40%	Dexter	Dexter Elementary	55%	53%
Central Consolidated	Nizhoni Elementary	29%	23%	Dora	Dora Elementary	45%	64%
Central Consolidated	Ojo Amarillo Elementary	31%	26%	Dulce	Dulce Elementary	33%	25%
Central Consolidated	Ruth N Bond Elementary	44%	35%	Elida	Elida Elementary		43%
Chama Valley	Chama Elementary	63%	58%	Espanola	Abiquiu Elementary	86%	65%
Chama Valley	Tierra Amarilla Elementary	59%	47%	Espanola	Alcalde Elementary	68%	48%
Cimarron	Cimarron Elementary	43%	50%	Espanola	Carinos De Los Ninos Charter	52%	47%
Cimarron	Eagle Nest Elementary	63%	82%	Espanola	Chimayo Elementary	29%	21%
Clayton	Alvis Elementary	82%	66%	Espanola	Dixon Elementary	85%	85%
Cloudcroft	Cloudcroft Elementary	66%	59%	Espanola	Eutimio Salazar Elementary	42%	35%
Clovis	Barry Elementary	47%	68%	Espanola	Hernandez Elementary	32%	23%
Clovis	Bella Vista Elementary	48%	30%	Espanola	James Rodriguez Elementary	54%	56%
Clovis	Cameo Elementary	44%	58%	Espanola	Mountain View Elementary		27%
Clovis	Highland Elementary	48%	57%	Espanola	San Juan Elementary	69%	60%
Clovis	James Bickley Elementary	52%	46%	Espanola	Tony Quintana Elementary	33%	32%
Clovis	La Casita Elementary	45%	52%	Espanola	Velarde Elementary	43%	56%
Clovis	Lockwood Elementary	47%	42%	Estancia	Estancia Upper Elementary	49%	60%
Clovis	Mesa Elementary	80%	75%	Eunice	Mettie Jordan Elementary	26%	44%
Clovis	Parkview Elementary	41%	32%	Farmington	Animas Elementary	53%	48%
Clovis	Ranchvale Elementary	79%	79%	Farmington	Apache Elementary	29%	41%
Clovis	Sandia Elementary	47%	57%	Farmington	Bluffview Elementary	36%	48%
Clovis	Zia Elementary	80%	78%	Farmington	Country Club Elementary	77%	82%
Cobre Consolidated	Bayard Elementary	65%	58%	Farmington	Esperanza Elementary	43%	64%
Cobre Consolidated	Central Elementary	54%	60%	Farmington	Ladera Del Norte Elementary	71%	61%
Cobre Consolidated	Hurley Elementary	80%	86%	Farmington	McCormick Elementary	30%	44%
Cobre Consolidated	San Lorenzo Elementary	60%		Farmington	McKinley Elementary	65%	69%
Corona	Corona Elementary			Farmington	Mesa Verde Elementary	56%	56%

District	School	SY11	SY12	District	School	SY11	SY12
Farmington	Northeast Elementary	57%	58%	Grants Cibola	Mount Taylor Elementary	48%	38%
Floyd	Floyd Elementary	79%	64%	Grants Cibola	San Rafael Elementary		
Fort Sumner	Fort Sumner Elementary	46%	70%	Grants Cibola	Seboyeta Elementary		
Gadsden	Anthony Elementary	66%	67%	Hagerman	Hagerman Elementary	53%	45%
Gadsden	Berino Elementary	45%	37%	Hatch Valley	Garfield Elementary	78%	50%
Gadsden	Chaparral Elementary	29%	33%	Hatch Valley	Rio Grande Elementary	37%	41%
Gadsden	Desert Trails Elementary	39%	38%	Hobbs	Broadmoor Elementary	72%	51%
Gadsden	Desert View Elementary	60%	42%	Hobbs	College Lane Elementary	62%	40%
Gadsden	Gadsden Elementary	61%	58%	Hobbs	Coronado Elementary	27%	27%
Gadsden	La Union Elementary	30%	51%	Hobbs	Edison Elementary	55%	55%
Gadsden	Loma Linda Elementary	44%	51%	Hobbs	Jefferson Elementary	31%	46%
Gadsden	Mesquite Elementary	30%	30%	Hobbs	Mills Elementary	58%	49%
Gadsden	North Valley Elementary	60%	49%	Hobbs	Sanger Elementary	65%	56%
Gadsden	Riverside Elementary	32%	37%	Hobbs	Southern Heights Elementary	31%	20%
Gadsden	Santa Teresa Elementary	67%	66%	Hobbs	Stone Elementary	71%	74%
Gadsden	Sunland Park Elementary	51%	60%	Hobbs	Taylor Elementary	36%	41%
Gadsden	Sunrise Elementary	51%	56%	Hobbs	Will Rogers Elementary	28%	25%
Gadsden	Vado Elementary	50%	38%	Hondo Valley	Hondo Elementary	9%	29%
Gallup McKinley	Chee Dodge Elementary	23%	28%	Jal	Jal Elementary	48%	63%
Gallup McKinley	Church Rock Elementary	36%	21%	Jemez Mountain	Gallina Elementary	44%	
Gallup McKinley	Crownpoint Elementary	23%	12%	Jemez Mountain	Lindrith Area Heritage Charter		
Gallup McKinley	David Skeet Elementary	6%	25%	Jemez Mountain	Lybrook Elementary	18%	8%
Gallup McKinley	Indian Hills Elementary	50%	44%	Jemez Valley	Jemez Valley Elementary	43%	29%
Gallup McKinley	Jefferson Elementary	34%	41%	Jemez Valley	San Diego Riverside Charter	27%	
Gallup McKinley	Juan De Onate Elementary	36%	35%	Lake Arthur	Lake Arthur Elementary		36%
Gallup McKinley	Lincoln Elementary	34%	31%	Las Cruces	Alameda Elementary	39%	34%
Gallup McKinley	Navajo Elementary	9%	27%	Las Cruces	Booker T Washington Elementary	48%	46%
Gallup McKinley	Ramah Elementary	35%	38%	Las Cruces	Central Elementary	42%	40%
Gallup McKinley	Red Rock Elementary	62%	57%	Las Cruces	Columbia Elementary	32%	36%
Gallup McKinley	Rocky View Elementary	16%	17%	Las Cruces	Conlee Elementary	34%	51%
Gallup McKinley	Roosevelt Elementary	63%	59%	Las Cruces	Desert Hills Elementary	70%	70%
Gallup McKinley	Stagecoach Elementary	20%	22%	Las Cruces	Dona Ana Elementary	49%	36%
Gallup McKinley	Thoreau Elementary	65%	39%	Las Cruces	East Picacho Elementary	55%	55%
Gallup McKinley	Tobe Turpen Elementary	27%	5%	Las Cruces	Fairacres Elementary	62%	53%
Gallup McKinley	Tohatchi Elementary	48%	38%	Las Cruces	Hermosa Heights Elementary	31%	43%
Gallup McKinley	Twin Lakes Elementary	25%	33%	Las Cruces	Highland Elementary	50%	64%
Gallup McKinley	Washington Elementary	28%	42%	Las Cruces	Hillrise Elementary	71%	65%
Grady	Grady Elementary			Las Cruces	Jornada Elementary	48%	47%
Grants Cibola	Bluewater Elementary	87%	53%	Las Cruces	Loma Heights Elementary	48%	50%
Grants Cibola	Cubero Elementary	41%	56%	Las Cruces	Mac Arthur Elementary	39%	61%
Grants Cibola	Mesa View Elementary	42%	39%	Las Cruces	Mesilla Elementary	63%	63%
Grants Cibola	Milan Elementary	43%	45%	Las Cruces	Mesilla Park Elementary	55%	54%

District	School	SY11	SY12	District	School	SY11	SY12
Las Cruces	Monte Vista Elementary	72%	59%	Mora	Mora Elementary	57%	65%
Las Cruces	Sonoma Elementary	77%	70%	Moriarty-Edgewood	Edgewood Elementary	85%	62%
Las Cruces	Sunrise Elementary	41%	44%	Moriarty-Edgewood	Moriarty Elementary	38%	45%
Las Cruces	Tombaugh Elementary	55%	49%	Moriarty-Edgewood	Mountainview Elementary	56%	57%
Las Cruces	University Hills Elementary	48%	66%	Moriarty-Edgewood	Route 66 Elementary	63%	84%
Las Cruces	Valley View Elementary	46%	60%	Moriarty-Edgewood	South Mountain Elementary	75%	78%
Las Cruces	White Sands Elementary	65%	65%	Mosquero	Mosquero Elementary		
Las Vegas City	Legion Park Elementary	41%	52%	Mountainair	Mountainair Elementary	56%	41%
Las Vegas City	Los Ninos Elementary	25%	76%	Pecos	Pecos Elementary	58%	46%
Las Vegas City	Mike Sena Elementary			Penasco	Penasco Elementary	54%	63%
Las Vegas City	Paul D Henry Elementary	67%	52%	Pojoaque Valley	Pablo Roybal Elementary	59%	63%
Las Vegas City	Sierra Vista Elementary	33%	43%	Portales	Valencia Elementary	57%	52%
Logan	Logan Elementary	82%	86%	Quemado	Datil Elementary		
Lordsburg	Southside Elementary	43%	44%	Quemado	Quemado Elementary		73%
Los Alamos	Aspen Elementary	73%	69%	Questa	Alta Vista Elementary	53%	27%
Los Alamos	Barranca Mesa Elementary	81%	83%	Questa	Rio Costilla Elementary		
Los Alamos	Chamisa Elementary	73%	74%	Raton	Columbian Elementary	71%	55%
Los Alamos	Mountain Elementary	91%	88%	Reserve	Glenwood Elementary		
Los Alamos	Pinon Elementary	76%	83%	Reserve	Reserve Elementary	72%	
Los Lunas	Ann Parish Elementary	40%	45%	Rio Rancho	Cielo Azul Elementary	59%	67%
Los Lunas	Bosque Farms Elementary	70%	68%	Rio Rancho	Colinas Del Norte Elementary	60%	55%
Los Lunas	Daniel Fernandez Elementary	35%	48%	Rio Rancho	Enchanted Hills Elementary	75%	75%
Los Lunas	Desert View Intermediate	42%	31%	Rio Rancho	Ernest Stapleton Elementary	64%	68%
Los Lunas	Katherine Gallegos Elementary	47%	53%	Rio Rancho	Maggie Cordova Elementary	76%	70%
Los Lunas	Los Lunas Elementary	48%	48%	Rio Rancho	Martin Luther King Jr Elementary	78%	68%
Los Lunas	Los Lunas Family School			Rio Rancho	Puesta Del Sol Elementary	51%	57%
Los Lunas	Peralta Elementary	48%	59%	Rio Rancho	Rio Rancho Elementary	67%	58%
Los Lunas	Raymond Gabaldon Elementary	53%	44%	Rio Rancho	Sandia Vista Elementary	61%	75%
Los Lunas	Sundance Elementary	61%	77%	Rio Rancho	Vista Grande Elementary	72%	66%
Los Lunas	Tome Elementary	51%	51%	Roswell	Berrendo Elementary	71%	73%
Los Lunas	Valencia Elementary	77%	59%	Roswell	Del Norte Elementary	56%	63%
Loving	Loving Elementary	35%	30%	Roswell	East Grand Plains Elementary	60%	56%
Lovington	Jefferson Elementary	58%	56%	Roswell	El Capitan Elementary	56%	32%
Magdalena	Magdalena Elementary	55%	12%	Roswell	Military Heights Elementary	72%	63%
Maxwell	Maxwell Elementary		70%	Roswell	Missouri Avenue Elementary	81%	42%
Melrose	Melrose Elementary	69%	38%	Roswell	Monterrey Elementary	49%	44%
Mesa Vista	El Rito Elementary	36%		Roswell	Nancy Lopez Elementary	34%	62%
Mesa Vista	Ojo Caliente Elementary		20%	Roswell	Pecos Elementary	69%	69%

District	School	SY11	SY12	District	School	SY11	SY12
Roswell	Sunset Elementary	41%	25%	Socorro	Parkview Elementary	59%	50%
Roswell	Valley View Elementary	50%	43%	Socorro	San Antonio Elementary		64%
Roswell	Washington Avenue Elementary	74%	47%	Springer	Wilferth Elementary	50%	50%
Roy	Roy Elementary			T or C	Arrey Elementary	10%	35%
Ruidoso	White Mountain Elementary	55%	42%	T or C	T Or C Elementary	55%	41%
San Jon	San Jon Elementary		67%	Taos	Anansi Charter	100%	85%
Santa Fe	Acequia Madre Elementary	68%	70%	Taos	Arroyo Del Norte Elementary	55%	56%
Santa Fe	Agua Fria Elementary	28%	38%	Taos	Enos Garcia Elementary	45%	43%
Santa Fe	Amy Biehl Community School at Rancho Viejo	69%	80%	Taos	Ranchos De Taos Elementary	35%	46%
Santa Fe	Aspen Community Magnet School	54%	44%	Taos	Taos Municipal Charter	92%	61%
Santa Fe	Atalaya Elementary	63%	73%	Tatum	Tatum Elementary	27%	40%
Santa Fe	Carlos Gilbert Elementary	65%	69%	Texico	Texico Elementary	73%	72%
Santa Fe	Cesar Chavez Elementary	38%	45%	Tucumcari	Tucumcari Elementary	59%	51%
Santa Fe	Chaparral Elementary	51%	56%	Tularosa	Tularosa Intermediate	50%	39%
Santa Fe	E J Martinez Elementary	59%	59%	Vaughn	Vaughn Elementary		
Santa Fe	El Dorado Elementary	74%	85%	Wagon Mound	Wagon Mound Elementary		
Santa Fe	Francis X Nava Elementary	19%	33%	West Las Vegas	Don Cecilio Martinez Elementary	44%	66%
Santa Fe	Gonzales Elementary	68%	69%	West Las Vegas	Rio Gallinas Ecology and the Arts Charter		36%
Santa Fe	Kearny Elementary	48%	48%	West Las Vegas	Tony Serna Jr Elementary	17%	47%
Santa Fe	Pinon Elementary	58%	61%	West Las Vegas	Union Elementary	58%	64%
Santa Fe	R M Sweeney Elementary	18%	37%	West Las Vegas	Valley Elementary	43%	50%
Santa Fe	Ramirez Thomas Elementary	18%	43%	Zuni	A:Shiwi Elementary	36%	
Santa Fe	Salazar Elementary	42%	28%	Zuni	Dowa Yalanne Elementary	53%	42%
Santa Fe	Santa Fe School For The Arts			State Charter	Albuquerque School of Excellence Charter	74%	68%
Santa Fe	Tesuque Elementary	19%	39%	State Charter	Albuquerque Sign Language Academy Charter		
Santa Fe	Turquoise Trail Elementary Charter	56%	52%	State Charter	Cien Aguas International		58%
Santa Fe	Wood Gormley Elementary	80%	84%	State Charter	Horizon Academy West	66%	55%
Santa Fe	Zia Behavior Class			State Charter	International School At Mesa Del Sol Charter		62%
Santa Rosa	Rita A Marquez Elementary	55%	9%	State Charter	J Paul Taylor Academy	89%	70%
Santa Rosa	Santa Rosa Elementary	78%	49%	State Charter	La Promesa Early Learning Center Charter	0%	14%
Silver Cons.	Cliff Elementary	64%	71%	State Charter	Montessori Elementary Charter	80%	74%
Silver Cons.	G W Stout Elementary	63%	71%	State Charter	New Mexico School For The Deaf		
Silver Cons.	Harrison Schmitt Elementary	74%	70%	State Charter	North Valley Academy	56%	54%
Silver Cons.	Jose Barrios Elementary	71%	47%	State Charter	Ralph J Bunche Academy		50%
Silver Cons.	Sixth Street Elementary	55%	60%	State Charter	Red River Valley Charter		
Socorro	Cottonwood Valley Charter	67%	40%	State Charter	Taos Integrated School for the Arts Charter	47%	81%
Socorro	Midway Elementary	31%	60%				

Source: PED

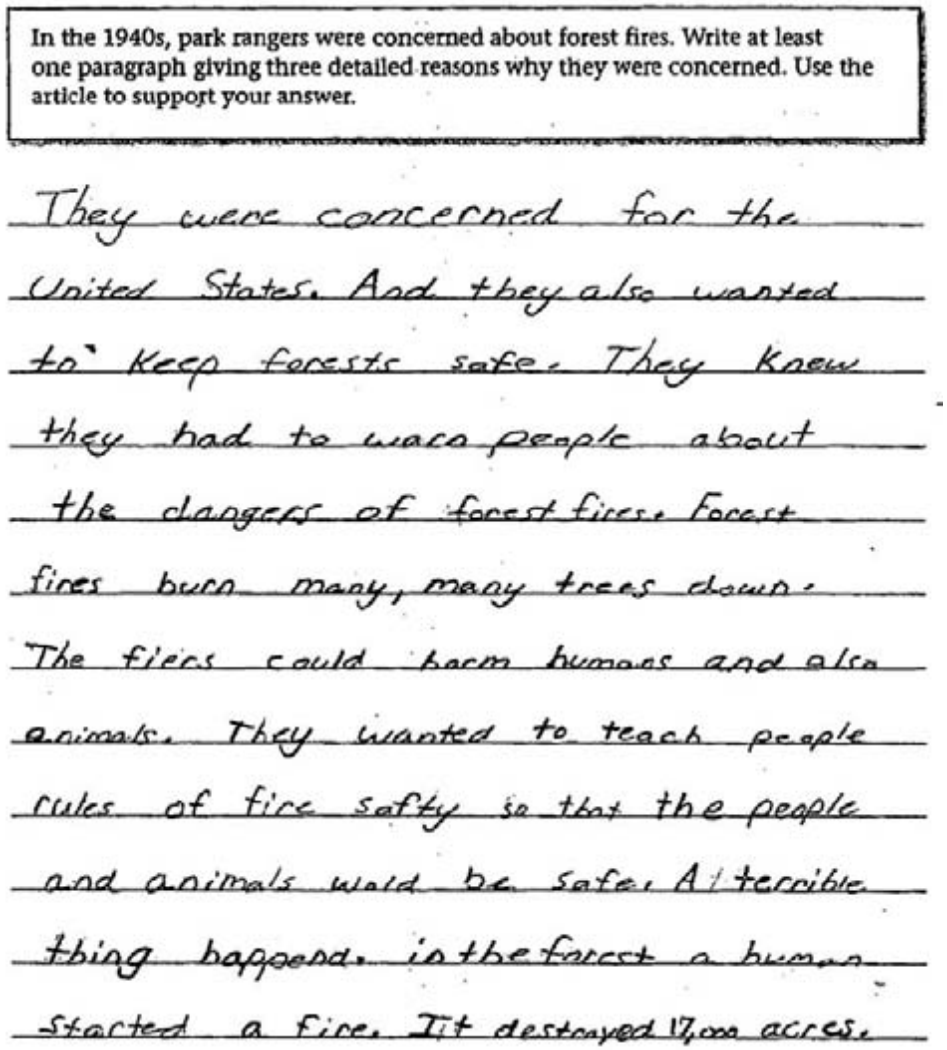
APPENDIX D: READING GRADE 3 SBA SAMPLE ITEM, SCORING GUIDE, AND STUDENT WORK

Item: In the 1940s, park rangers were concerned about forest fires. Write at least one paragraph giving three detailed reasons why they were concerned. Use the article to support your answer.

Scoring Rubric:

Score	Description
Four points	Response gives three detailed reasons why they were concerned.
Three points	Response gives two detailed reasons why they were concerned.
Two points	Response gives one detailed reasons why they were concerned.
One point	Response states that they were concerned.
Zero points	Response is totally inappropriate and includes irrelevant details

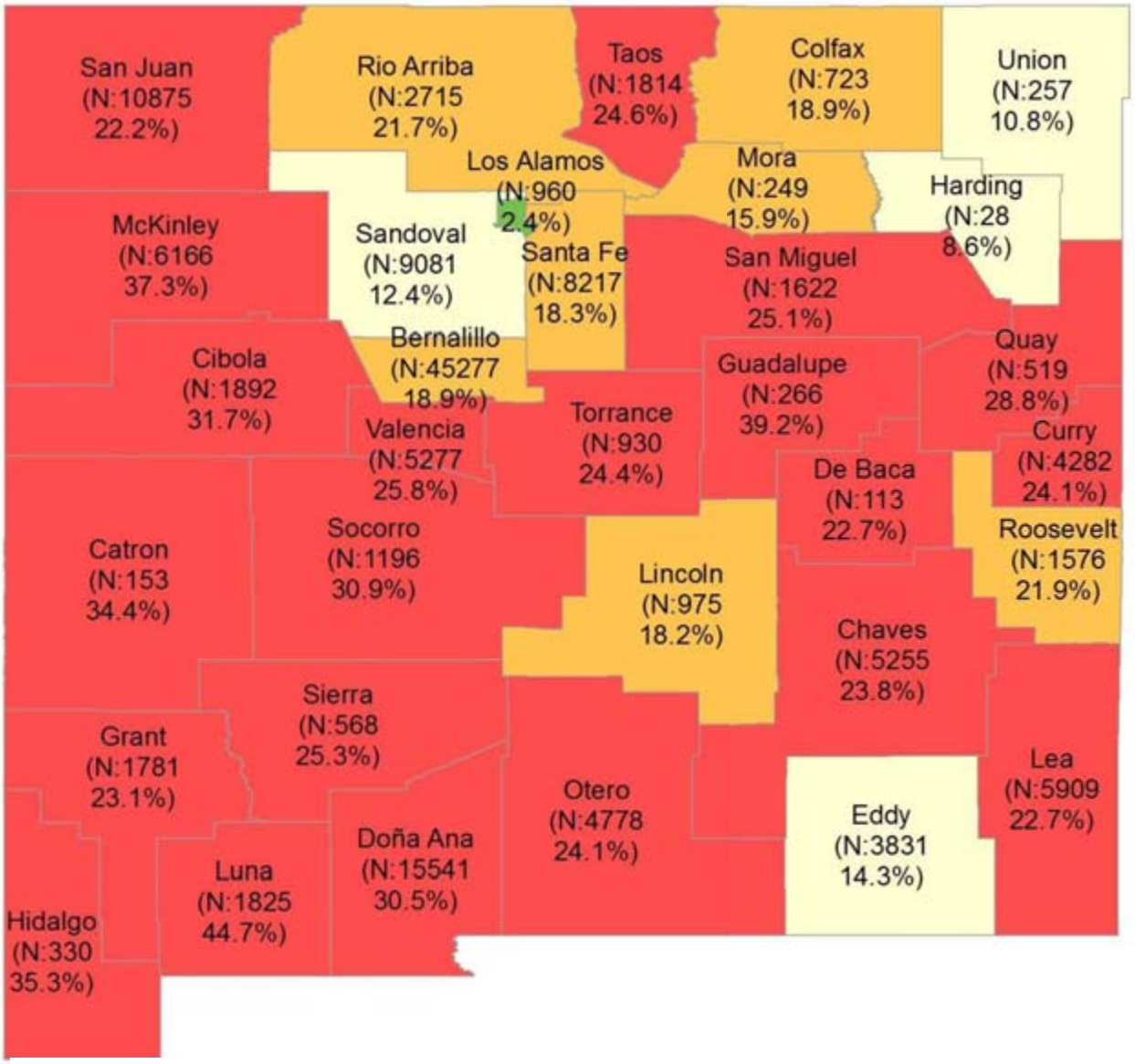
Four Point Response:



In the 1940s, park rangers were concerned about forest fires. Write at least one paragraph giving three detailed reasons why they were concerned. Use the article to support your answer.

They were concerned for the United States. And they also wanted to keep forests safe. They knew they had to warn people about the dangers of forest fires. Forest fires burn many, many trees down. The fires could harm humans and also animals. They wanted to teach people rules of fire safty so that the people and animals wold be safe. A terrible thing happend. is the forest a human started a fire. It destroyed 17,000 acres.

APPENDIX E: PERCENTAGE OF FAMILIES WITH CHILDREN WITH INCOME LESS THAN 100 PERCENT OF THE FEDERAL POVERTY LEVEL



Source: Center for Educational Policy Research

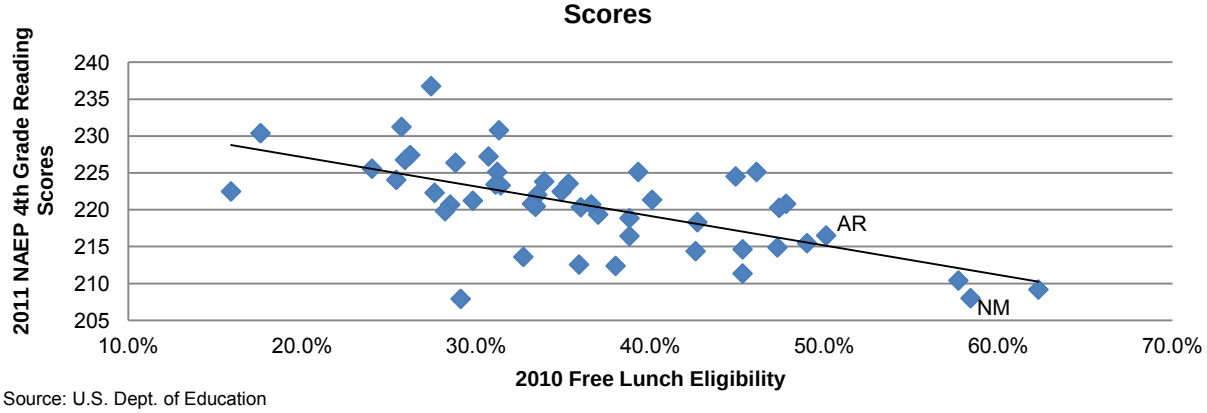
Even when controlling for poverty, New Mexico's students lag behind the nation in reading assessment scores. States with high rates of students eligible for free or reduced-price lunch (FRL) generally have lower reading scores on the National Assessment of Educational Progress (NAEP). States with up to 30 percent of FRL students averaged a NAEP reading score of 225, while states with over 50 percent of FRL students averaged a score of 211. As poverty rates increase, average state reading scores on the NAEP decrease.

Table 22. State FRL Eligibility and NAEP Reading Scores

FRL Range	Average 4th grade NAEP Reading Score	% lower than FRL Eligibility of 15%-30%
0% - 30%	225	-
30% - 40%	221	-2%
40% - 50%	218	-3%
50% and above	211	-6%

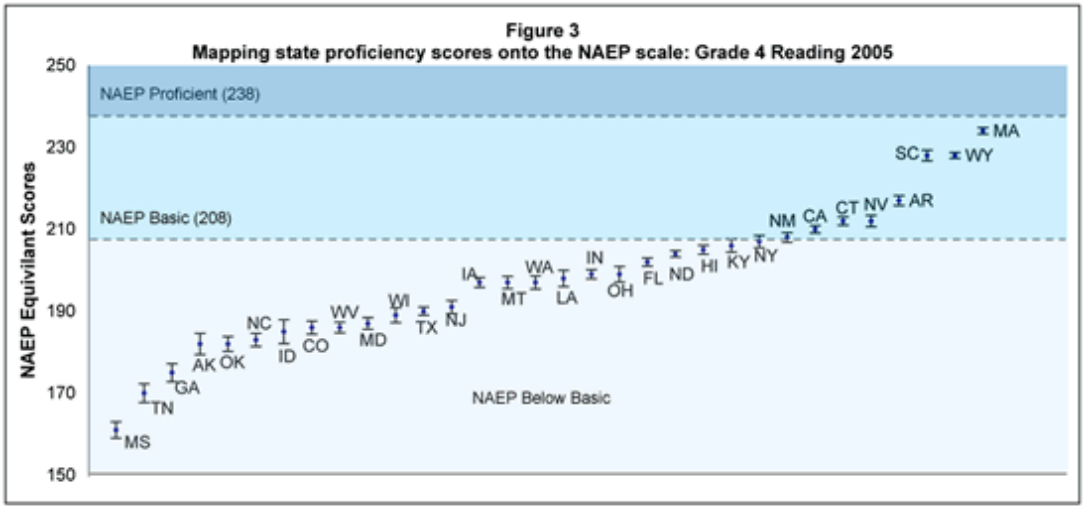
New Mexico performs worse than expected given its student demographics. Based on 2011 NAEP scores and FRL rates for all states, New Mexico's NAEP score of 208 is three points below its predicted value of 211. New Mexico's fourth-grade reading scores underperform the most among high-poverty states. Only one state with over 50 percent of FRL students, Arkansas, out-performed this predicted value.

Chart 13. State Free Lunch Eligibility Rates and NAEP 4th Grade Reading Scores



Source: U.S. Dept. of Education

There is a large gap between NAEP proficiency levels and state measures of proficiency, but that gap is relatively small in New Mexico, suggesting a relatively rigorous state standards-based test. According to a 2005 study conducted by the Education Trust, NAEP's proficiency rates for fourth-grade reading are on average 40 percentage points less than proficiency rates of state assessments. As measured by NAEP, 21 percent of New Mexico fourth-graders are proficient or advanced in reading, while the SBA shows 53 percent of third-graders were proficient or advanced in reading in 2011. The difference between the two tests' proficiency rates, 32 percent, is not as large as many other states. In 2007, the National Center for Education Statistics statistically linked each state's 2005 standards-based test scores onto the NAEP's scoring scale. As shown below, New Mexico's proficient achievement level is equal to the NAEPs basic level.



Source: NCES (2007) Mapping 2005 State Proficiency Standards Onto the NAEP Scales (NCES 2007-482)

This holds true today, as 53 percent of New Mexico students scored basic and above on the NAEP and 53 percent scored proficient and above on the third-grade SBA reading assessment in 2011.

Table 23. Percentage of Students in Each SBA and NAEP Proficiency/Achievement Category, SY11

SBA Proficiency Level	Percent of Students	NAEP Achievement Level	Percent of Students
Advanced	6%	Advanced	3%
Proficient	47%	Proficient	17%
Nearing Proficiency	25%	Basic	33%
Beginning Steps	22%	Below Basic	47%

NAEP confirms that large gaps related to eligibility for free or reduced-price lunch. Since 2005, FRL students have scored 12 percent below non-FRL students on the fourth-grade NAEP reading assessment. These factors have an impact on the differences in NAEP scores between students of different ethnic backgrounds. For example, 83 percent of Hispanic students and 90 percent of Native American students qualify for FRL, compared with 44 percent of Caucasian students. These rates corresponded with their NAEP fourth-grade reading scores. In 2011, Caucasian students scored an average of 225 on the assessment, while Hispanic and Native American students scored an average of 199 and 190, respectively.

APPENDIX G: SCHOOL PROFILES FOR 12 SITE VISITS

District & School	Average Number of third-Graders	Percent Proficient on SY06 third-grade Reading SBA	Percent Proficient on SY11 third-grade Reading SBA	FRL	ELL	Hispanic	Native American
Albuquerque Public Schools	7,196	54%	52%	62%	27%	58%	5%
Dolores Gonzales	74	41%	51%	99%	42%	93%	2%
Emerson	74	14%	4%	100%	51%	69%	6%
Griegos	56	72%	78%	63%	6%	76%	2%
Belen Consolidated Schools	357	55%	54%	76%	18%	71%	2%
Jaramillo	92	52%	55%	100%	11%	70%	2%
Central Consolidated Schools	462	40%	41%	100%	35%	2%	89%
Newcomb	55	24%	61%	100%	65%	0%	98%
Ojo Amarillo	50	59%	31%	100%	73%	0%	100%
Espanola Public Schools	369	40%	51%	70%	61%	90%	7%
Hernandez	38	63%	32%	100%	70%	96%	0%
Las Cruces Public Schools	1,843	57%	52%	65%	25%	72%	1%
Conlee	91	56%	34%	74%	27%	84%	0%
Mesilla Park	87	51%	55%	77%	39%	75%	2%
Roswell Independent School District	751	53%	60%	73%	14%	65%	0%
Sunset	51	49%	41%	100%	30%	80%	0%
Washington Ave	73	59%	74%	100%	10%	62%	1%
Santa Fe Public Schools	1,068	49%	50%	67%	38%	75%	3%
Chaparral	63	52%	51%	24%	17%	67%	4%
Statewide	24,931	55%	53%	66%	15%	57%	11%

Source: PED

APPENDIX H: PROCEDURE TO ESTIMATE PREK EFFECT

1. Used Pearson correlation to determine which variables have a statistically significant relationship with CYFD and PED PreK:

Correlations

		Hispanic	Caucasian	Native American	Other	ELL	Sped	SBA SCORE	CYFD PreK SY07	PED PreK SY07	FRL
Hispanic	Pearson Correlation	1	-.725**	-.409**	-.269**	.212**	-.045**	-.148**	.043**	-.014*	.286**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.027	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
Caucasian	Pearson Correlation	-.725**	1	-.184**	-.121**	-.270**	.053**	.234**	-.033**	-.066**	-.377**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
Native American	Pearson Correlation	-.409**	-.184**	1	-.068**	.084**	-.008	-.120**	-.016*	.139**	.120**
	Sig. (2-tailed)	.000	.000		.000	.000	.216	.000	.010	.000	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
Other	Pearson Correlation	-.269**	-.121**	-.068**	1	-.055**	.007	.032**	-.009	-.027**	-.057**
	Sig. (2-tailed)	.000	.000	.000		.000	.255	.000	.146	.000	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
ELL	Pearson Correlation	.212**	-.270**	.084**	-.055**	1	-.002	-.320**	-.027**	.053**	.247**
	Sig. (2-tailed)	.000	.000	.000	.000		.787	.000	.000	.000	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800

Sped	Pearson Correlation	-.045**	.053**	-.008	.007	-.002	1	-.201**	-.025**	-.026**	-.029**
	Sig. (2-tailed)	.000	.000	.216	.255	.787		.000	.000	.000	.000
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
SBA SCORE	Pearson Correlation	-.148**	.234**	-.120**	.032**	-.320**	-.201**	1	.014*	-.010	-.289**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.024	.107	.000
	N	24800	24800	24800	24800	24800	24800	25495	25495	25495	24800
CYFD PreK SY07	Pearson Correlation	.043**	-.033**	-.016*	-.009	-.027**	-.025**	.014*	1	-.032**	.012
	Sig. (2-tailed)	.000	.000	.010	.146	.000	.000	.024		.000	.060
	N	24800	24800	24800	24800	24800	24800	25495	25495	25495	24800
PED PreK SY07	Pearson Correlation	-.014*	-.066**	.139**	-.027**	.053**	-.026**	-.010	-.032**	1	.075**
	Sig. (2-tailed)	.027	.000	.000	.000	.000	.000	.107	.000		.000
	N	24800	24800	24800	24800	24800	24800	25495	25495	25495	24800
FRL	Pearson Correlation	.286**	-.377**	.120**	-.057**	.247**	-.029**	-.289**	.012	.075**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.060	.000	
	N	24800	24800	24800	24800	24800	24800	24800	24800	24800	24800
**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed)											

2. Identified students who had attended CYFD PreK, FY07 built an ANCOVA with the following co-variates: ethnicity, Sped, ELL, FRL.

   Output:

2. CYFD PreK SY07

Dependent Variable: UPDATED SBA SCORE

CYFD PreK SY07	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
0	39.537[a]	.067	39.405	39.669
1	39.999[a]	.377	39.259	40.738

a. Covariates appearing in the model are evaluated at the following values: Hispanic = .62, Caucasian = .25, Native American = .09, Other = .04, ELL = .21, Sped = .16, FRL = .7359.

3. Estimated the effect of CYFD PreK at 0.4 scaled score points.
4. Repeated the same ANCOVA for PED PreK:

2. PED PreK SY07

Dependent Variable: UPDATED SBA SCORE

PED PreK SY07	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
0	39.492[a]	.067	39.360	39.624
1	41.301[a]	.371	40.574	42.027

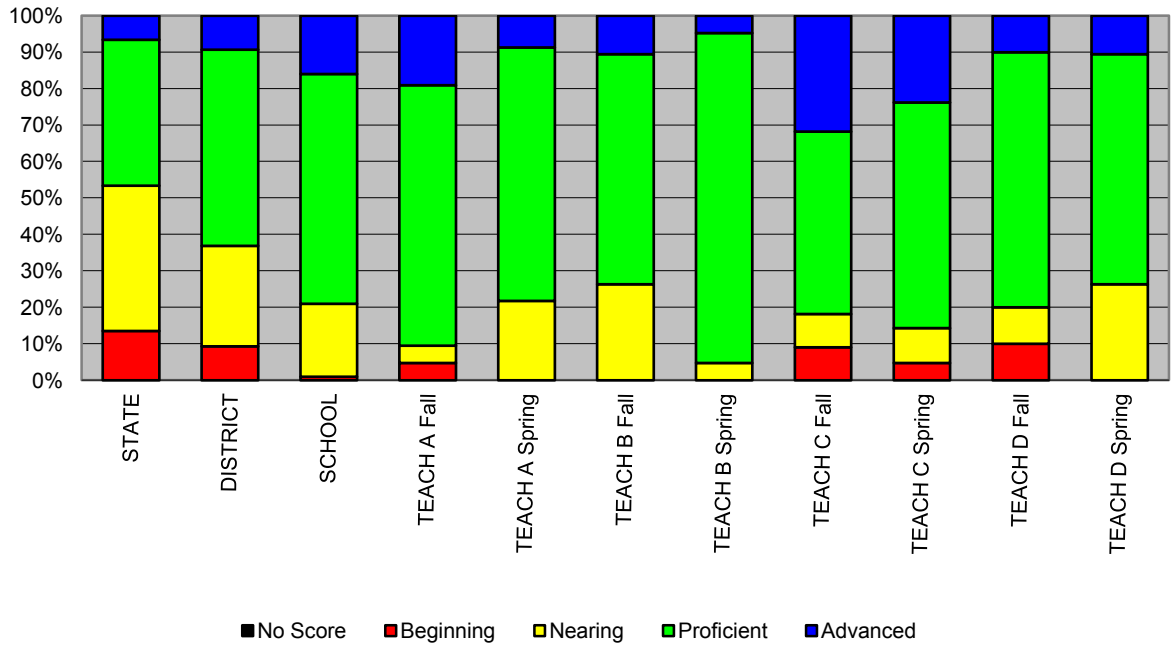
a. Covariates appearing in the model are evaluated at the following values: Hispanic = .62, Caucasian = .25, Native American = .09, Other = .04, ELL = .21, Sped = .16, FRL = .7359.

5. Estimated the effect of PED PreK at 1.8 scaled score points.

APPENDIX I: HARRISON SCHMITT ELEMENTARY DATA EXAMPLE

Harrison Schmitt Elementary in Silver City uses data to track student performance by classroom from the beginning of the school year to the end.

Table 24. School Year Classroom Performance by Teacher, SY11



Source: Silver City Consolidated School District