



School Success Summit

Measuring Student Learning
and Teacher Effectiveness

FINAL REPORT

- This document summarizes community input regarding the possibility of adopting growth models as evaluation tools.
- Summit held April 26, 2011, Albuquerque Convention Center

CONVENER

Albuquerque Public Schools

FACILITATOR

New Mexico First



NEW MEXICO FIRST

People. Ideas. Progress.

Copyright 2011

New Mexico First
P. O. Box 56549
Albuquerque, New Mexico 87187
Phone: 505-241-4813
Website: www.nmfirst.org

New Mexico First typically grants permission to reproduce all or part of this document for educational or public policy purposes. Contact us for reproduction authorization.

CONTENTS

CONVENER	1
Contents	3
Introduction and Overview	5
Convener	5
Facilitator	5
Summary of Results	6
Key Idea #1: A SYSTEM Better Than AYP?.....	6
Key Idea #2: Objective Tool for Assessment.....	6
Key Idea #3: Potential Negative Affects on Teaching	7
Key idea #4: Potential for Tailored Instruction	7
Key Idea #5: Concern About the “Whole Student”	7
Key Idea #6: Concern about More Testing	8
Key Idea #7: Caution-Go Slow	8
Key Idea #8: Adequate Funding.....	9
Appendix A	10
Definitions of Evaluation Models.....	10
Status Models.....	10
Improvement Model (type of status model)	10
Growth Model	10
Value-Added Models (type of growth model)	10
Appendix B	12
Top Priorities for Each Group.....	12
QUESTION 1: What is appealing or intriguing about growth models?.....	12
Community and Business Group	12
District and Administration Executives.....	12
Parent Group	12
Teacher Group.....	13
Principal Group.....	13
Research Groups (two groups)	13
Superintendent Group.....	13
QUESTION 2: What concerns exist about growth models?	14

Community and Business Group14

District and Administration Executives14

Parent Group14

Teacher Group.....14

Principal Group.....14

Research Groups15

Superintendent Group.....15

QUESTION 3: Questions district or state leaders must address before pursuing a growth model15

Community and Business Group15

District and Administration Executives15

Parent Group16

Teacher Group.....16

Principal Group.....16

Research Groups16

Superintendents Group16

INTRODUCTION AND OVERVIEW

In this age of school accountability, everyone wants to know that America’s children are receiving quality educations. How do we know if students are developing the skills they need for a bright future? Do current evaluation models, such as standardized tests, give us the information to judge successful schools and effective teaching? To what degree should our school evaluation models include teacher effectiveness? Throughout the nation, school districts and states are experimenting with different ways to answer these questions.

On April 26, 2011, Albuquerque Public Schools joined in this national conversation by convening a dialogue to introduce “growth models” and discuss the pros and cons. Growth models are evaluation approaches that measure student achievement over time, as opposed to “status models” (which measure student achievement at a given point in time, such as Adequate Yearly Progress).

About 150 people took part in the School Success Summit, learning about various evaluation models from leading experts. These national speakers addressed how other school districts and states use growth and status models. Participants broke into stakeholder groups and offered their own impressions, questions, and concerns about the possible integration of growth models into APS’ school evaluation processes. Groups included:

- Parents
- Teachers
- Principals
- District administrators
- Superintendents
- Researchers
- Business people
- Community members
- APS board members
- Legislators
- Public Education Department employees

Convener

Albuquerque Public Schools serves 90,000 students, about one third of all New Mexico students. The district is comprised of 142 elementary, middle, and high schools and 65 departments.

One of those departments, Instruction and Accountability, initiated the School Success Summit summarized in this report. The mission of Instruction and Accountability is to create a culture of support for instructional improvement. Essential to the culture is a commitment to essential standards, aligned curriculum, effective classroom instruction for students, ongoing assessment underlying data-driven instructional decision making, and evaluations of efficacy of practice and program.

Facilitator

New Mexico First engages people in important issues facing their state or community. Established in 1986, the public policy organization offers unique town halls and forums that bring together people from all walks of life to develop their best ideas for policymakers and the public. New Mexico First also produces nonpartisan public policy reports on critical issues facing the state. These reports – on topics like water, education, healthcare, the economy, and energy – are available at nmfirst.org.

Our state’s two U.S. Senators – Jeff Bingaman and Tom Udall – serve as New Mexico First’s honorary co-chairs. The organization was cofounded in 1986 by Senators Jeff Bingaman and Pete Domenici (retired).

SUMMARY OF RESULTS

The following summary identifies crosscutting ideas that emerged from multiple stakeholder group discussions. Participants frequently used the terms “growth model” and “value-added model” interchangeably but without necessarily understanding the difference. Throughout this report, the term “growth model” is used for consistency.

KEY IDEA #1: A SYSTEM BETTER THAN AYP?

Each of the stakeholder groups recognized the positive potential for the use of growth models over the current AYP status model.

Many participants expressed the need to move away from the notion of proficiency toward a “more fair” measurement-based system, one that would give credit to schools that make progress, versus the arbitrary distinctions inherent in the current Annual Yearly Progress. The concept of using growth models could be validating for so-called “failing schools,” since it is not punitive but rather proactive, celebrating the growth of the student and building on the positive accomplishments. Many participants believed that a growth model could render a fairer evaluation for schools that may not make the AYP status cut.

Unlike present standardized testing measures, developing a growth model could provide districts and schools with the opportunity to develop measures that are useful to teachers, students and the community. Not only do growth models move away from a “snap shot” toward a “bigger picture” of achievement by focusing on assessment beyond one point in time, they also use multiple sources of data to measure the value added by a school or teacher rather than just one piece of data. Participants appreciated the possibility of a school community coming together to define and measure what it values, see what is missing, and determine how to measure growth over time.

KEY IDEA #2: OBJECTIVE TOOL FOR ASSESSMENT

Stakeholders also saw the potential for growth models to provide a more objective way to assess teachers and principals and provide feedback for professional development.

Stakeholders emphasized the importance of identifying excellent teachers beyond test scores and expanding our means of assessing teacher effectiveness. A growth model could recognize the inspirational and role modeling aspect of teaching. It could help verify what teachers are doing correctly, and validate teachers who perform effectively, particularly with lower performing students. Teacher and principal evaluations could also become more valuable, providing more opportunity for professional growth and development.

Participants emphasized that teachers’ voices should always be a part of the process in developing any new growth model. It would be important to involve teachers in developing assessment measures and to explore how effective teachers evaluate themselves. These non-punitive models could help new and evolving teachers become aware of how they can reflectively self-evaluate and improve. In helping determine what good teachers do and how to replicate it, participants believed a growth model could provide feedback to teachers that would lead to modified instructional activities.

KEY IDEA #3: POTENTIAL NEGATIVE AFFECTS ON TEACHING

At the same time, there were concerns about the potential negative effects of the use of a growth model on teachers and teaching.

There was concern that a growth model would allow scapegoating of teachers when a problem is system-wide. Said one participant, “Teacher evaluation is like a pie, and value-added assessment is only a piece of the pie, but I am concerned that it will be used as the whole pie.” Some expressed fears that using growth models for teacher evaluation and compensation could be punitive. One asked, “Is this a move to streamline and reduce the teaching workforce? Will these models negatively impact teacher morale?”

Several groups also mentioned that if growth models were used to determine compensation or merit pay, competition among teachers could hinder current attempts to foster collaboration. “If we attach compensation to a growth model, it could pit teachers against each other. How would I, as a principal, navigate pay for performance?”

KEY IDEA #4: POTENTIAL FOR TAILORED INSTRUCTION

Some stakeholders saw the potential for a growth model to positively affect students, better assess student progress, and tailor instruction for individual students.

Some participants believed that a growth model would enable individual data and information to follow each student, creating a better picture of individual student progress. It would provide information to students, parents, teachers that is useful for students with disabilities, and enable teachers to adjust instruction accordingly. For some, growth models “celebrate student success—and make us step back and examine how we evaluate student knowledge. It helps us break the cycle of perceived failure.”

KEY IDEA #5: CONCERN ABOUT THE “WHOLE STUDENT”

Participants were concerned about a growth model not valuing and supporting the “whole student” and addressing the diversity of individual students.

On this matter, participants had more questions than answers. Several asked some variation of the question, “How do we define growth so that it takes into consideration the emotional and social development and growth of students?”

Participants wondered how to measure student growth for students who have family or other issues outside of the school that may not be measurable. They were concerned about how to measure children and youth who test poorly, such as special education students and English language learners. Commented one participant, “We can measure students against students, but students are not equal. What kind of student do we want to produce?”

KEY IDEA #6: CONCERN ABOUT MORE TESTING

Stakeholder groups expressed concern about the complexity of developing appropriate standards for evaluation and uses of data, as well as the inevitability of having to spend even more time on testing and data collection than they now do.

Common concerns included: whether future evaluation models would ask the right questions; whether the education community could come to consensus on a purpose and get new data to support that purpose; and what would be done with the data once it is gathered. Participants also raised concerns about how to support teachers while measuring their outcomes. Several voiced concerns about aligning current professional development efforts to a new system.

Said one attendee, “The fact that you get new numbers doesn’t tell you what to do about professional development. If we identify not very good teachers and we pile new responsibilities on to existing ones, how does that relate to current professional development plans?”

Concerns were also expressed about current models of measurement as well as the possibilities for these new growth models. For example, “Can these models assess higher level thinking and the ability to synthesize information?” It was pointed out that the focus is very narrow with current models. “The measures are way too limited,” said one attendee. “Why are we only testing math and reading? What about creative problem solving, music, and being a good human being?”

There was also concern that adopting growth models would result in more testing, taking time away from instruction. Some participants were concerned that increased testing would further strain limited time available for subjects like recess, art, music, AP and honors classes. Said one, “It’s going to be a fine line and so much information that we may drown in data. How much worse is it going to get than it is already, with the testing we have now?”

KEY IDEA #7: CAUTION-GO SLOW

The fear of poor implementation was a recurring concern for stakeholders, who repeatedly said to “go slowly,” take adequate time for thoughtful and reflective planning, and “don’t rush.”

Stakeholders repeatedly emphasized concerns about implementing a system that has not been completely been thought out, that would not reliably measure outcomes and that would lose its focus on student growth. Putting a policy in place before the process was thought out, or before stakeholders had a chance to weigh in, would be a mistake, participants believed. Said one, “I fear a ‘rush job’ by state officials and knee-jerk reactions to the state of our schools. I caution them to go slowly and listen to teachers on the front lines.”

Strategies for addressing this concern included: developing an overall plan with timelines; considering teacher concerns; developing common set of core standards; and taking adequate time to roll out a new system. Some participants believed it was important to avoid creating fear in school employees. Asked one participant, “At the state level, are we willing to do it right, at a slower pace, with no short cuts?”

One group said that the pace would be determined by the answer to the question, “Is this process politically driven or educationally driven?” Several groups also suggested testing any new models with a pilot process,

before implementing a growth model statewide. Said one participant, “People might look for a silver bullet instead of gathering all the needed data to answer the questions – you can’t just run the model and expect student scores to improve.”

KEY IDEA #8: ADEQUATE FUNDING

Stakeholders emphasized the importance of ensuring adequate funding and other resources necessary to effectively implement and sustain a growth model.

An adequate budget would be required to do a thorough job collecting multiple measures and creating a more robust picture. It would be important to help the community understand that a significant amount of money would be required to develop a more rigorous form of assessment.

There were concerns that the public and policymakers may not understand all of the time, money and work needed. “You need to get that right when the stakes are high. There is a record of not fully funding things.” It was felt that it was important to prioritize resources and determine the sources of funding.

Appendix A

DEFINITIONS OF EVALUATION MODELS

SOURCE: Excerpted from the Council of Chief State School Officers' Policymakers' Guide to Growth Models for School Accountability. The full report is available on their website (www.ccsso.org).

Status Models

Status models are often contrasted with growth models. A status model (such as Adequate Yearly Progress, or "AYP") takes a snapshot of a subgroup's or school's level of student proficiency at one point in time and often compares that proficiency level with an established target. Progress can be defined, for example, as the percentage of students achieving at the proficient level for that particular year, and the school is evaluated based on whether the student group met or did not meet the goal.

Improvement Model (type of status model)

A status model analyzes school educational achievement compared against an established performance target -- usually for one specific year. In addition, status models can be compared at two points in time to provide a measure of improvement. An improvement model is a type of status model that measures change between different groups of students (i.e., the performance of this year's fourth graders compared with last year's fourth graders).

Growth Model

Growth models generally refer to models of educational accountability that measure progress by tracking the achievement scores of the same students from one year to the next with the intent of determining whether or not, on average, the students made progress. For example, learning growth can be measured by comparing the performance of this year's fourth graders with the performance of the same students last year in third grade. Achievement growth over time at the school level is then the aggregate of growth data for individual students, controlling for each students' background and prior achievement. By comparing data for the same students over time, progress can be defined as the degree to which students' estimated improvement compares to a statewide or local target. Growth models assume that student performance, and by extension school performance, is not simply a matter of where the school is at any single point in time. A school's ability to facilitate academic progress is a better indicator of its performance. Growth models vary greatly.

Value-Added Models (type of growth model)

A commonly referenced application of a growth model is the value-added model. VAMs are an approach in which states or districts use student background characteristics and/or prior achievement and other data as statistical controls in order to isolate the specific effects of a particular school, program, or teacher on student academic progress. The main purpose of VAMs is to separate the effects of non-school factors (such as family, peer, and individual influence) from a school's performance at any point in time, so that student performance can be attributed appropriately. A value-added estimate for a school is simply the difference between its actual

growth and its expected growth. It is important to note that schools can demonstrate positive achievement growth, but have a value-added estimate this is negative (i.e., the school demonstrated growth, just not as much as would have been predicted given the student inputs available to the school). A well known type of VAM is the Tennessee Value-Added Assessment System. Like most growth models, it tracks the yearly growth in student learning. However, this model measures student growth by modeling a series of gains in performance demonstrated by each student as well as the teachers who instructed and the schools that provided the context for their instruction. Thus, the model attempts to attribute the change in performance of students to the specific providers of instruction during a specific period.

Appendix B

TOP PRIORITIES FOR EACH GROUP

Each participant joined a discussion group during the summit, based on her/his stakeholder group. Each group developed responses to each of the following questions, and then prioritized them. These priorities informed the previous section (thus are repeated in some cases). They are listed here by group, so that readers can see how different stakeholders perspectives vary.

QUESTION 1:

What is appealing or intriguing about growth models?

COMMUNITY AND BUSINESS GROUP

- Growth models move past the notion that standardized testing is the best or only way to assess quality teaching. Growth models can expand our means of assessment of teacher effectiveness. Growth models move assessment beyond one point in time and take into account the diversity of student backgrounds.
- Growth models can celebrate student success, making us step back and examine how we evaluate student knowledge. They help break the cycle of perceived failure.
- The involvement of teachers in determining the assessment structure is important. These models allow for input by teachers. We can usefully examine how teachers evaluate themselves. These models can help new and developing teachers to become aware of how they can self-evaluate. These models can assist the large number of “average” teachers to self-reflect and improve their skills as educators.

DISTRICT AND ADMINISTRATION EXECUTIVES

- Pace of the rollout.
- Are the models political or educational?
- Sustainability and funding.

PARENT GROUP

- We like that the district would hold teachers and districts more accountable for students learning. We want to find a way to hold teachers accountable and measure student growth.
- Parents are excited for new information that follows each student. Less lump sum and more individual data is a good way to go; parents want a foot in the door to see how kids are doing.
- Parents find the measures interesting, but want to know how to help and how the growth model will incorporate parents. Following up at home on the foundation laid at school is important.
- The questions the summit did not ask were intriguing to parents. Kids are dying in many ways (incarceration, drugs, etc). With that going on for decades, why are we having a summit on growth models vs. value-added models? When are we going to have a summit on why are our kids dying? Said one: “We’re rearranging the furniture on the Titanic.”

TEACHER GROUP

- Educators may be a part of the process that makes the system work.
- Growth models provide a method for the use of multiple sources of data and urges us to use caution while assessing.
- Growth models offer the potential to develop measures that are more useful to teachers, students, and community than are the current assessments.

PRINCIPAL GROUP

- The notion of deeply examining what it is that we (APS, or “my school”) value, how we measure that, and coming together within the school to see what is missing is intriguing. Also interesting is the ability to define and measure what we value as a school community and how measure growth over time.
- With reference to Senate Bill 502, which did not pass but dealt with teacher and principal evaluation, this discussion helped define value-added models.
- We would like to see this evolve into something that affects student success. If done well, it could help us improve in a number of areas, and help identify what makes an effective teacher and principal.

RESEARCH GROUPS (TWO GROUPS)

Intrigued by the notions of:

- Data over time (longitudinal vs. point-in-time).
- Rendering a fair evaluation for schools that may not make the AYP status cut.
- The opportunity to determine the effects of different factors and tease out quality teaching measures.
- Setting aside the status model of NCLB, to offer more valuable, useful, reasonable information about how students are doing to teachers, media, and the public.
- Setting up the model so that you can really do something with it (it is only valuable if it helps us improve).
- Starting by asking reasonable and important questions about what good teaching is and how to identify it.
- When creating your own system to meet school, district and state needs and purposes, models can be open-ended and open to all kinds of data.
- Understanding goals and purposes, recalling that growth models are only valuable if used to help students improve.

SUPERINTENDENT GROUP

- Growth models give credit to schools that make progress, versus arbitrary distinctions from AYP.
- Growth models provide a better picture of individual student progress.
- Growth models validate effective teachers, especially with lower performing students (such as special education).

QUESTION 2:

What concerns exist about growth models?

COMMUNITY AND BUSINESS GROUP

- What kind of student growth is being measured and which teachers are evaluated?
- Can these models assess higher-level thinking and the ability to synthesize information?
- Do these models address concerns to the public?
- There is a concern about moving quickly to tie assessment to compensation. Go slow. Will these models negatively affect teacher morale?
- How do these models address truancy? Why are students not coming to school? Why are students not learning the basics? How do these models address early learning?

DISTRICT AND ADMINISTRATION EXECUTIVES

- At the state level, are we willing to do it right – slower pace, no short cuts?
- Is this the right place to put our efforts? Are we doing the right work? Are we asking the right question(s)?
- Will we plan a pilot to learn unintended consequences?

PARENT GROUP

- Growth models can allow scapegoating--in this case, the teachers--when the problem is system-wide. We want to support leadership of teachers, schools, students and parents.
- What system can we put into place to be sure we are not chasing a red herring? We are getting the “whats,” but not the “hows,” nor the results of the implementation.
- What are the negative impacts of the data-collecting emphasis on our students? More testing time? Less time or resources for things like recess, art, music, AP and honors classes.

TEACHER GROUP

- Who makes the decisions and whose voices will be heard in the development of this model? Teachers' voices in developing model should always be a part of the process.
- Teachers are worried about implementing a system that has not been completely thought through and that does not reliably measure outcomes. They fear a “rush job” by state officials and “knee-jerk” reactions to the state of our schools. Teachers caution to go slowly and listen to educators on the front lines. If not, the focus on student growth will be “lost in the rush.”
- The reality is that assessment will be used for evaluation and compensation, but should not.

PRINCIPAL GROUP

- Concerned about rushing through the process. (Concerned about a policy being put in place before the process is thought out or stakeholders have comprehensive involvement). Need a pilot project first.
- If we attach compensation to a growth model, it could pit teachers against each other. How would principals navigate pay for performance (competitiveness among staff members, schools)?
- We need teachers, parents, administrators, board members to be educated on growth models before moving forward.

RESEARCH GROUPS

- Will a growth model help us continue to develop a culture of collaboration between teachers or will the growth model foster a culture of competitiveness between teachers?
- Currently we only have three subjects measured (reading, math, and science). For an initial model, this is satisfactory as long as the model has a vision for more subjects.
- A percentile model creates a distribution of high vs. low teachers. We need the absolute measures included in addition to the percentile measures.
- An adequate budget is required to do a good job collecting multiple measures and creating a more robust picture; people might look for a silver bullet instead of gathering all needed data to answer questions; don't expect to just run the model and see student scores improve.
- Make sure there is the time and money to do growth modeling correctly. There is a concern that state lawmakers will need to be convinced to do this right.
- Growth models present major communication challenges for parents and the community.
- It is critical to take time for thoughtful and reflective planning; don't take the first "draft" we see.

SUPERINTENDENT GROUP

- Have we identified the purpose of the growth model system?
- Make sure it provides a better picture of individual student progress.
- A concern is the ambiguity about the state's impending evaluation system requirements; will it be a broad framework or an imposed system?

QUESTION 3:**Questions district or state leaders must address before pursuing a growth model****COMMUNITY AND BUSINESS GROUP**

- What is the focus /purpose of wanting the information about student assessment and teacher effectiveness?
- Who has access to which data?
- How will you take into account the grades and subjects not covered in SBA?
- Who will be involved in the evaluation of teachers, particularly if evaluation dimensions are expanded to different measures?

DISTRICT AND ADMINISTRATION EXECUTIVES

- What do we do as a result of this growth model?
- How do we know the growth is a direct influence of the teacher?
- How will we help the community and stakeholders understand that it will take a significant amount of money to develop assessment of grades and content that is not currently evaluated rigorously?

PARENT GROUP

- What are schools looking for with growth models – what is the purpose? Where do these models come from – do they come from a successful school district?
- Where will we get money to implement these new models? What might we sacrifice to fund them?
- Why are schools systematically destroying children’s enthusiasm for learning in the pursuit of data and new data models?
- What are schools doing to address an “organizational culture” (how we do business) at the schools? Why are parents unwelcome?

TEACHER GROUP

- How will the data be used to implement changes in curriculum and instruction?
- We need to decide what is the goal of graduates? College? Workforce? Vote? Etc.
- What do community members and teachers need to know? What do growth models mean to me and how am I involved?

PRINCIPAL GROUP

- What is the definition of the growth model, and how will stakeholders understand this?
- How does this tie to a principal’s evaluation? How does the principal manage a growth model, and how is it reflected in the evaluation domains? How is it tied to student achievement?
- How do you measure growth at a school that performs in the higher or lower quartile?

RESEARCH GROUPS

- What is the purpose of the growth model? What is the goal of the growth model by stakeholder groups? Identify the stakeholders that need to be included in the conversation and the effective implementation of the growth model.
- Anticipate some of the outcomes in advance: Establish an organized, transparent, and reasonably well-planned method of introducing the growth model, gathering input, responding to concerns and questions, revising the model, communicating results and findings, etc.
- What resources are needed and how do we get them? Where will they come from?
- What role does deployment of Common Core Standards have as we think about deploying growth models?
- How would new models align with existing systems (e.g., Professional Development Plan system, tier system, etc.)
- How public will this data be?
- What implications exist for APS (or other school districts) if PED creates a separate or different growth model system?
- What resources will be deployed for educating people about the system, results, etc.?

SUPERINTENDENTS GROUP

- What assessments will be used to determine growth?
- How will growth model connect w/ the A-F grading system?
- How can the districts collaborate (pool resources) to facilitate state development of the new accountability system?