

Middle States Regional Forum, Using Research to Increase Teacher Efficacy: Outcomes from the 3-Year EE4NJ Pilot Confirmation

Dr. Michael Kuchar, Superintendent Bergenfield NJ

Dr. Frank Auriemma, EE4NJ Consultant, retired Superintendent Pearl River NY

Dr. Cynthia Blitz, Research Professor, Executive Director Center for Effective School Practices, Rutgers University

Ms. Elizabeth Jewett, Superintendent Watchung Regional HS District, and Educational Consultant

Crisis in America

Historical Context of the Problem

Randi Weingarten, American Federation of Teachers President



“...With rare exceptions, teacher evaluation procedures are broken, cursory, perfunctory, superficial and inconsistent.

The Chicago Teacher Strike

The Chicago teachers strike drags into a second week, after a representative group of the Chicago Teachers Union decided over the weekend not to end the walkout even though union leaders and school officials had reached a tentative contract deal. The strike in the third-largest school system in the country is affecting more than 350,000 children.

A quick primer:

Q. What's the sticking point?

A. Among the major issues, the teachers are negotiating over the length of the school day, objecting to their evaluations being tied to performance and fretting about potential job losses.

CNN (September 17, 2012)

Management's Motive in Chicago

In 2010 high school graduation rate for Chicago Public Schools was 55.8 percent.

During the same time period, only .03% of all Chicago Public School teachers received an unsatisfactory rating on annual evaluation.

Rethinking Teacher Evaluations: Findings from the Excellence in Teaching Project Chicago Public Schools (2010). Consortium on Chicago School Research by the Urban Education Institute, University of Chicago.

The Widget Effect

Fewer than one percent of the 40,000 teachers covered by the study had ever received a lower than “satisfactory” rating on an evaluation.

Effective teachers are the key to student success. Yet our school systems treat all teachers as interchangeable parts, not professionals. Excellence goes unrecognized and poor performance goes unaddressed. This indifference to performance disrespects teachers and gambles with students' lives.

The Teacher Project (2009)

The Outcome



The national high school graduation rate is reported to be between 72-75 percent, dependent upon the data source. Currently, one in four students drops out before he/she finishes high school. That's one every 26 seconds or more than one million students a year.

America's Promise Alliance (2013)

Linking Student Achievement to Teacher Practice

More than two decades of research findings are unequivocal about the critical connection between teacher effectiveness and student learning. The research shows that student achievement is strongly related to teacher quality; highly skilled teachers produce improved student results.

Excellent Educators for New Jersey (EE4NJ)

- Because our district was already working to increase student performance through teacher practice, and our evaluators were already working to become better observers through 5D Assessment and Instructional Rounds, the next logical step was to tie in new teacher performance measures.
- New Jersey is reforming teacher evaluations, and it is better to be providing insight and feedback to the state rather than to be forced into a new system with no input.

Data Sources

- Danielson teacher observation data
- Ren-Learn SGP Data
- DOE SGP Data
- University of Washington 5D Evaluator Training data

Analyzed by Center for Effective Teaching, Rutgers University

Dr. Auriemma

Will talk about our data sources and using district data to improve student learning which ultimately leading to higher student achievement.

Bergenfield's Initial Work and Study

- Although created as a tool to improve teacher practice...Bergenfield selected the Danielson Model as their teacher evaluation instrument and served as a New Jersey EE4NJ Pilot District.
- Greatest challenge: For teachers to focus on feedback and identified areas of growth rather than just the ratings.
- Through an initial double blind study using teacher volunteers, we were able to collect data which aligned with the University of Chicago Danielson study...relative to Rater Accuracy and Agreement along with Component Rigor.

Teacher Voice and Choice

- Over three years, 50 teacher leaders and administrators spearheaded the rollout of the Danielson Framework.
- Funded training and extensive committee work at significant costs
- We addressed thorny issues including:
 - weighting of Domains
 - incorporating student learning outcomes
 - and the use of short cycle assessments
- And while democracy is a tedious process...

“No one got hurt” and committee work, even when stressful, maintained professionalism and civility.

5D Assessment and Instructional Rounds

- 5D Assessment (U. of Wash.) was used as a primary administrator training and assessment tool, providing rater accuracy scores through the use of video exemplars.
- Administrators receive accuracy scores based on both local and nationally normed performance.
- Our initial district study utilized multiple evaluators observing the same lesson and provided administrators an opportunity to compare, discuss and reflect upon their scores.
- Administrators and teachers developed common language, lexicon and “Look Fors” which have become more sophisticated over time.

Use of Formative Assessments

- The Formative Assessment format allows for a school or district wide short cycle assessment pattern...however some teachers utilize a much shorter cycle for their individual classroom needs.
- Use of Formative Assessment SGPs provide another valuable, albeit different, measure of student skill based growth.
- Teacher, Principal and Superintendent Dashboards provide ease of monitoring and data collection.

Bergenfield: Analyses of Teacher
Observation and Student Growth Data

Center for Effective School
Practices,
Rutgers University

Cindy Blitz, Ph.D.

Data Sources

- Danielson Teacher Observation Data
- Ren-Learn SGP Data
- NJDOE SGP Data

Types of Analyses

- Observation scores by subject, grade level, and tenure status
- Ren-Learn and NJDOE SGP scores by subject, grade level, and tenure status
- Direct comparison of Ren-Learn SGP and NJDOE scores
- Correlation between SGP scores & observation scores

Findings from Observation Data

- Untested teachers tend to score higher in domains 2 and 3 (observed domains) and lower in domains 1 and 4 (planning and professional responsibilities)
 - Out of these untested subjects, Specials and Science receive the highest observation scores
- Teachers in elementary schools tend to get higher observation scores than those in middle or high schools
- Tenured teachers tend to score significantly higher on observation ratings than teachers who are untenured

Ren-Learn vs. NJDOE SGP

- Ren-Learn SGP
 - Growth relative to peers **nationwide** on REN-Learn Assessments which are linked to Common Core
 - Compares growth from baseline assessment in the fall to two time points: end of fall, end of spring. Also compares growth from end of fall to end of spring
- NJDOE SGP
 - Growth relative to peers **statewide** on NJASK
 - Compares growth from end of **previous** academic year to end of following academic year
 - Calculated for a smaller subset of teachers (20+ students, grades 4-8)
- What they have in common
 - Both compare a student's growth relative to a group of his peers. Thus, an SGP of 50 indicates that a student grew more than 50% of his peers **who scored similarly** on the baseline assessment

Summary of SGP Findings

- Median Ren-Learn SGP for teachers in Bergenfield is in the high-typical range, while NJDOE SGP is in the low typical range
- Teachers in elementary school classrooms tend to have higher growth in their classrooms, on average, compared to middle school teachers, but comparable to high school teachers
- Tenured teachers, much like in observation scores, receive higher median SGPs as well
- Regarding the discrepancy between Ren-Learn SGP and NJDOE SGP: even when we only look at the 34 teachers who had both SGPs, still significant differences

Summary and Synthesis of Findings

- Across domains, teachers tend to receive higher observation scores in elementary schools than in middle or high schools
- Across observation domains and SGPs, tenured teachers tend to perform better than untenured teachers
- Teacher observation ratings have a moderate correlation to Ren-Learn SGP
- Strong, positive correlations between Ren-Learn SGP & NJDOE SGP, but NJDOE scores generally lower

Implications and Questions to Consider

- What accounts for the difference in growth rate between elementary and high school levels (high-typical) and middle school (typical)?
- Why are observation ratings in high school lower than in elementary despite similar SGP outcomes?
- Would there be any benefits of having observers external to the building? District level?
- What are tenured teachers doing differently that is leading to both higher observation scores and higher student outcomes?
 - Is there any way to harness “expert” teachers to improve outcomes for novices?
- What sort of training *would* help administrators conduct evaluations more consistently?
- Why is there a decrease in student growth rate during the winter-spring semester?

Administrator responsibility

- School leaders must have a clear idea of what good instruction should look like. In addition, instructional leaders need to know why they make certain instructional decisions over others in order to have a lasting impact on teaching and learning.

Data: The “Heart” of the SGO?



Data from Rutgers, University of Washington, and student data to drive Professional Development

Goal for Today's Session

- Together we will collaboratively create a Comprehensive Professional Development program for administration and teachers that supports the Common Core/PARCC using the district strategic plan as a vehicle for student success and all available data points (research).

SUMMARY OF FINDING FROM THE OBSERVATION DATA (Rutgers University)

- Teachers who don't teach Math and English tend to score higher in domains 2 and 3 (observed domains) and lower in domains 1 and 4 – planning and professional responsibilities.
 - Out of these untested subjects, Specials and Science receive the highest observation scores.
- Teachers in elementary schools tend to get higher observation scores than those in middle or high schools.
 - *But, do those reflect real differences in teaching or evaluator bias?*
- Tenured teachers tend to score significantly higher on observation ratings than teachers who are untenured.

REVIEW OF EARLIER FINDINGS OF PILOT DATA

- On average, observations in Bergenfield tend to yield higher scores than those sampled in the MET study.
- On average, teachers tend to score lower in domain 3 – Professional responsibility than in the other 3 domains.
 - Specifically, teachers tend to score lower in 3b – Using questions and discussion techniques and 3d – Using assessment.
- Across measures of inter-rater agreement, evaluators in Bergenfield tend to show moderate to good agreement on ratings.
 - However, agreement does not necessarily equal accuracy.
 - We hope to investigate accuracy of ratings by comparing to SGP data.

Creating Your SGO Blueprint!



Just released!

<http://www.state.nj.us/education/AchieveNJ/teacher/SGOGuidebook.pdf>

Principal as Instructional Leader: Guiding the SGO Conversation



PREPARE SGO

- **Is there a viable curriculum in place that is standards-based and connected to daily instruction?**
- **What assessment data is available? What does the assessment data tell us about our students? Are additional assessments needed? If so, can these be procured or will teachers need to collaboratively create them?**
- **Are the assessments aligned to the standards?**

PREPARE SGO

- **Are the pre-assessments, formative assessments and post-assessments aligned?**
- **Are PLCs/teams established and working effectively to analyze student data and identify areas of student need?**

DEVELOP SGO

- **What student learning objectives will be the basis of my SGO?**
 - **What kind of SGO will I develop?**
 - **Progress and/or achievement?**
 - **General or General-tiered?**
 - **Specific – student group?**
 - **Specific- content/skill?**
- **If PLCs/teams have identified an area of student need, have individual teachers developed growth goals specifically for their students?**

DEVELOP SGO

- **Are PLCs/teams established and working effectively to identify the most effective teaching and learning strategies to maximize student achievement and meet the SGO?**
- **Do teachers and/or PLCs/teams need targeted professional development regarding specific strategies or the diverse needs of learners?**

IMPLEMENT and MONITOR the SGO

- **Is student assessment data shared among colleagues so that the needs of students can be more effectively met?**
- **Are PLCs/teams collaboratively reflecting on and revising teaching and learning strategies and instructional resources in response to analysis of assessment data?**

SGO's for Administrators



Instructional Rounds

- Assessing what we see in a classroom rather than what we feel in relation to the learning goals of the class, school and district.
- We have utilized Instructional Rounds as a learning vehicle for both teachers and administrators.
- By visiting classrooms in groups of three, we have been able to establish rater agreement at approximately 85%. Such agreement would be meaningless, except for the fact that we have already established rater accuracy via the 5D Assessment Model.
- Our challenge going forward is to annually assess administrators to assure continued and improving rater competencies.

EE4NJ Bergenfield Rater Reliability

- Using Domains 2 and 3, we now have 10 components; 2a to 2e (Organizing Physical Space) and 3a to 3e (Demonstrating Flexibility) as recommended by both Danielson and Teachscape*

Principal Evaluation: *SGP and SGO Components*

School SGP

- **Principals** whose students have SGPs **will receive the average school-wide SGP score.**
- **Principals will be placed in 3 categories:** Multi-Grade SGP Principal, Non-SGP Principal, Single-Grade SGP Principal. Component weighting will differ across categories.

SGO Average

- Principals will be rated on **their teachers' success in achieving student growth objectives** (SGOs) each year through an average of their teachers' scores.

		Components	Multi-Grade SGP Schools	Non-SGP Schools	Single Grade SGP Schools
Inputs		Principal Practice Instrument	30%	30%	30%
		Evaluation Leadership	20%	20%	20%
Student/ Teacher Outcomes		SGO Average	10%	10%	10%
		School SGP	30%	0%	20%
		Principal Goals	10%	40%	20%
		Total Percentage	100%	100%	100%

S.M.A.R.T SGO's are...

Specific	Measurable	Attainable/Ambitious	Results-driven	Timed
The SGO should be simplistically written, and clearly defined.	The SGO should be measurable and provide tangible evidence that you have achieved the objective.	The SGO should be attainable; reasonably challenging both you and your students, but clearly defined so that it can be achieved.	The SGO should focus on measuring outcomes, not activities.	The SGO should be organized around a timeframe that presents a reasonable sense of urgency.

TYPES OF SGOs

Type of SGO	Definition
General	Focused on the teacher's entire student population for a given course. Includes a large proportion of curriculum standards
General – Tiered	Same as above, but with student goals tiered by student preparation levels.
Specific – Student Group	Focused on a subgroup of students that needs specific support.
Student – Content/Skill	Focused on specific skills or content that students must master.

TYPE: General SGO – Elementary Literacy

SGO Statement:	80% of students increase at least one proficiency level on the Text Reading and Comprehension (TRC) assessment.			
Measuring Progress	For a teacher to earn a rating of...			
	4	3	2	1
	*90% of more students met goal.	*80% or more students met goal.	*70 or more students me their goal	*Less than70% of students me their goal

*These numbers will be determined by teacher and principal based on knowledge of students to create a rigorous and attainable goal

TYPE: Tiered General SGO – Physics 1

SGO Statement:	75% students will meet their designated target scores on the Physics 1 post assessment.
-----------------------	--

Preparedness Group	No. of Students in Group	Target Score on PA (%)
Low	36/65	70
Medium	21/65	80
High	8/65	90

Measuring Progress	For a teacher to earn a rating of...			
	4	3	2	1
Low	85% or more students in the tier met goal.	75% or more students in the tier met goal.	65% or more students in tier met goal	Less than 65% of students in tier met goal
Medium				
High				

TYPE: Specific/Targeted Students – Gr. 8 ELA

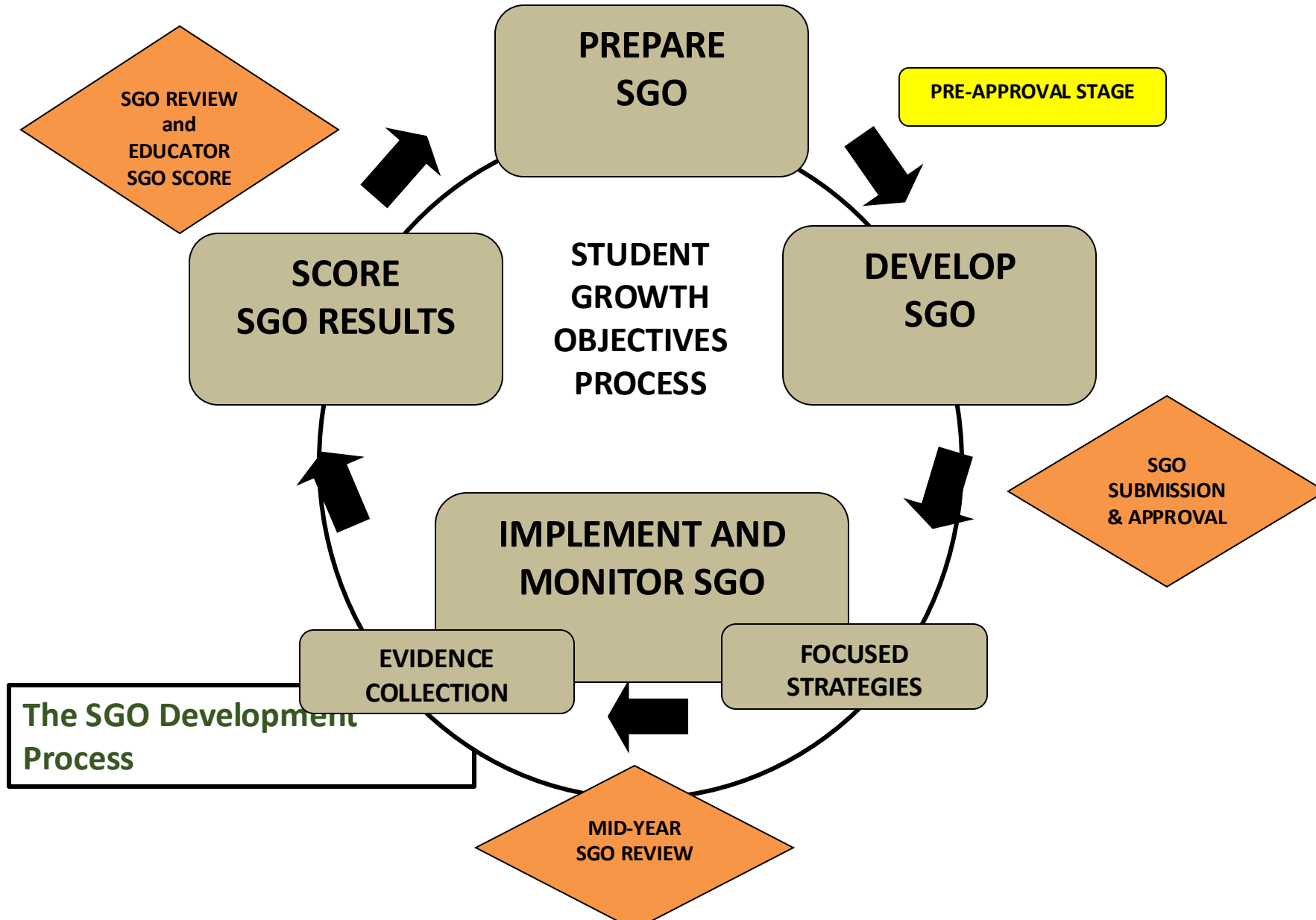
SGO Statement:	6/8 students who scored in the low range on the pre-assessment will increase 10 words/minute over their baseline score on the Oral Reading Fluency Assessment.			
Measuring Progress	For a teacher to earn a rating of...			
	4	3	2	1
	7-8 students met goal	5-6 students met goal.	3-4 students met goal	0-2 students met goal.

For some teachers there may be a specific student group that is appropriate to target. In this instance, the teacher identified a group of students with low preparedness who he believed would benefit from increased work in reading fluency.

TYPE: Specific/Targeted Content/Skill- History

SGO Statement:	80% of students will score a “3” or better on the district DBQ assessment for using evidence to support a point of view.			
Measuring Progress	For a teacher to earn a rating of...			
	4	3	2	1
	90% or more students met goal.	80% or more students met goal.	70 or more students met their goal	Less than 70% of students met their goal

Teachers can also use rubrics or portfolio assessments to measure student attainment. In this example the district created a rubric for U.S. History students to measure attainment of specific critical thinking skills.



Parking Lot/To Do from 6/23

1. Align Report Cards to CC/Assessments Utilized
2. Revisit the domains and rubrics in Danielson
3. Conversion chart - Ren/SGP (Calibrate score)
4. How do we account for lack of writing/open ended response in Ren?
5. PD opportunities for teacher to be a part of a best practice model
6. iPad for admins w/Danielson software
7. Unit planning – UbD – curriculum maps?
8. PD – Eval – Common Core – all areas / teacher strategies, - CC, PARCC
9. PD – Teacher/Admin shared training
10. Danielson - calibration among raters, accuracy, agreement

- What actions have been taken regarding implementation of the CCSS at each level?
- What work remains to be done?
 - Teacher
 - Administration
 - Students
 - Parents/Community

- What would a district look like that is fully aligned to CCSS/PARCC?

Essential Question for 2014-2015...

How do we make Professional Development an integral component of the Strategic Plan in order to drive student achievement?

Some samples of Bergenfield's Success

Questions