Accountability in the Age of Pearson: Advocating for What works in Schools

Robert and Augusta P. Finkelstein Memorial Lecture
Adelphi University
Ruth S. Ammon School of Education

30 October 2013

Sean C. Feeney, Ph.D.
Principal
The Wheatley School
Old Westbury, NY 11568

www.newyorkprincipals.org / @scfeeney
Spackenkill HS Town Hall

www.newyorkprincipals.org / @scfeeney
Accountability in Schools

www.newyorkprincipals.org / @scfeeney
NYS APPR Legislation

- Passed by Legislature in May 2010
- Required student test scores to be used in evaluating teachers and principals
- NY SED Guidelines distributed in April 2011
- Effective in September 2011

www.newyork principals.org / @scfeeney
New York Principals
APPR Position Paper

www.newyorkprincipals.org / @scfeeney
Concern 1: Research does not support it

1. Value-added models (VAM) of teacher effectiveness do not produce stable ratings of teachers
   a) Measurement errors
   b) Inter-temporal stability concerns
   c) Year to year correlations for elementary school teachers are in the 0.2-0.4 range

2. No evidence that evaluation systems incorporating student test scores produce gains in student achievement

3. NYS exams are designed to evaluate student learning, not teacher effectiveness or student learning growth

www.newyork principals.org / @scfeeney
Concern 2: Students will be adversely affected

1) Narrowing of curriculum to focus on end of year examinations
2) System will be incentivize to place struggling students in “safer” classes
3) Teachers will be incentivized to avoid students with issues — health issues, ELL, emotional
4) Dynamic between students and teacher will change. “Teacher and student versus the exam” is replaced by “teacher versus students’ performance”
5) Collaboration among teachers replaced by competition
Concern 3: Costs to districts excessive

1) 10 days of APPR training for building principals; 17 days for Superintendents

2) Tax dollars redirected to testing companies for correction, security and analysis of exams
What have we seen since APPR was instituted?

We have become Test Crazy!

www.newyorkprincipals.org / @scfeeney
Student Testing

Commissioner John King:

- writes “..the number and frequency of state assessments has remained relatively constant over the last decade…”

- and explains that “…a maximum of 540 of those [64,800 instructional] minutes are devoted to State tests in Grades 3-8 ELA and math…”

- while boasting of a decrease in testing in 2013 from 2012

(October 24, 2013 “Teaching is the Core” memo)

www.newyorkprincipals.org / @scfeeney
NYS Testing Minutes in Grades 3-8 (Math and ELA)

Yes...that’s 9 hours of testing!

www.newyorkprincipals.org / @scfeeney
# NYS Testing Minutes in Grades 3-8 (Math and ELA)

<table>
<thead>
<tr>
<th></th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 6</th>
<th>Grade 7</th>
<th>Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>160</td>
<td>300</td>
<td>165</td>
<td>255</td>
<td>190</td>
<td>320</td>
</tr>
<tr>
<td>2009</td>
<td>160</td>
<td>300</td>
<td>165</td>
<td>265</td>
<td>190</td>
<td>320</td>
</tr>
<tr>
<td>2010</td>
<td>160</td>
<td>300</td>
<td>170</td>
<td>265</td>
<td>200</td>
<td>320</td>
</tr>
<tr>
<td>2011</td>
<td>250</td>
<td>300</td>
<td>260</td>
<td>290</td>
<td>295</td>
<td>305</td>
</tr>
<tr>
<td>2012</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
</tr>
<tr>
<td>2013</td>
<td>420</td>
<td>440</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
</tr>
<tr>
<td>% Change</td>
<td>163%</td>
<td>47%</td>
<td>227%</td>
<td>112%</td>
<td>184%</td>
<td>69%</td>
</tr>
</tbody>
</table>

[www.newyorkprincipals.org](http://www.newyorkprincipals.org) / @scfeeney
# 9 Hours of NYS ELA and math Testing!

<table>
<thead>
<tr>
<th>Test</th>
<th>When taken</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP exams</td>
<td>Grades 10-12, for college credit</td>
<td>3 hours</td>
</tr>
<tr>
<td>SAT exam</td>
<td>Grades 11-12, for college admission</td>
<td>3.3 hours</td>
</tr>
<tr>
<td>ACT exam</td>
<td>Grades 11-12, for college admission</td>
<td>3.5 hours</td>
</tr>
<tr>
<td>GRE</td>
<td>for graduate school admissions</td>
<td>3.75 hours</td>
</tr>
<tr>
<td>LSAT</td>
<td>for law school admissions</td>
<td>3.5 hours</td>
</tr>
<tr>
<td>MCAT</td>
<td>for medical school admissions</td>
<td>4.3 hours</td>
</tr>
<tr>
<td>CPA</td>
<td>to become a certified public accountant</td>
<td>14 hours</td>
</tr>
</tbody>
</table>

[www.newyorkprincipals.org](http://www.newyorkprincipals.org) / @scfeeney
But there are more tests...
SLO tests

- In order to evaluate teachers based on student test scores, more tests are needed!
- Teachers are administering exams based on Student Learning Objectives (SLOs)
- ALL subjects without a NYS test must have one of these SLOs
  - PE, art, technology, family and consumer science, kindergarten
- The pre-test / post-test comparison will be used to evaluate teachers
Field tests

- Just after completing two weeks of testing (for Math and ELA), 3,600 schools were required to administer field tests to these same students.
- SED is very quiet about the administration of these tests.
- The results of the field tests are used to help Pearson fulfill the terms of their $32 million NYS contract.
Lots and lots of tests...
What are we doing with all these tests?
Evaluating teachers

- 25% to 40% of a teacher’s end of year evaluation will depend on how his or her students performed on these tests
- Passing is no longer sufficient; students must show growth every year
Evaluating principals

- 25% to 40% of a principal’s end of year evaluation depend on how students perform on these tests
Improving schools?

Washington, DC Performance Index (under IMPACT)

Pearson Stock Price, 2009-present

www.newyorkprincipals.org / @scfeeney
We are losing our way

- Dylan was in second grade
- One of five children
- Mom completed 7th grade

www.newyorkprincipals.org / @scfeeney
Dylan’s Reading Teacher

- Dylan’s report card

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>GRADING PERIOD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>READING</td>
<td>C</td>
</tr>
<tr>
<td>LANGUAGE</td>
<td>A</td>
</tr>
<tr>
<td>SOCIAL STUDIES</td>
<td>S</td>
</tr>
<tr>
<td>SCIENCE/HEALTH</td>
<td>S</td>
</tr>
<tr>
<td>MATHEMATICS</td>
<td>B</td>
</tr>
</tbody>
</table>

www.newyorkprincipals.org / @scfeeney
Dylan’s Reading Test

- Florida Assessment for Instruction in Reading (FAIR), a computerized assessment
- After three tests in second grade:
  - Accuracy: 95%
  - Vocabulary percentile rank: 88th
  - Spelling percentile rank: 19th
- Determined that he would need to work on multisyllabic words
- Probability of Reading Success (PRS) score of 44% (probability of reading at or above grade level)
Test trumps Teacher

- Because he did not “pass” the FAIR test, Dylan was retained in 2nd Grade
  - Despite teacher comments regarding Dylan’s progress and achievement as a reader
  - Despite the overwhelming evidence that grade retention simply does not help students “catch up”
Test trumps Reality

- Sept 2009: Dylan’s grandfather dies
- Feb 2010: Dylan’s father is killed
- Jan 2012: Dylan’s Grandmother dies of cancer
- June 2012: Dylan does not “pass” the FAIR test
What are standardized tests telling us about schools and students?
http://nces.ed.gov/surveys/pisa/
http://nasspblogs.org/principaldifference/2010/12/pisa_its_poverty_not_stupid_1.html
Poverty and US Schools — 2011

http://www.southerneducation.org/Programs/P-12-Program/Early-Ed/NewMajority.aspx

www.newyork principals.org / @scfeeney
Percentage of low-income students

2000

Note: The report did not include D.C. because its school system is too small relative to those of states.

2011

*students who qualify for free or reduced-priced lunches


http://www.washingtonpost.com/local/education/poverty-increasing-in-public-schools/2013/10/16/94defd76-36d9-11e3-80c6-7e6dd8d22d8f_graphic.html

www.newyorkprincipals.org / @scdeeney
Teachers who receive the state’s top value-added rating — "Most Effective" — are likely to be in schools with fewer poor students, based on value-added ratings for teachers at 1,720 public schools. Of 1,035 teachers at the wealthiest schools, 34 percent got the top rating. In contrast, of 2,411 teachers at the poorest schools, just over 9 percent were rated "Most Effective."

Teachers who receive top value-added ratings are more likely to be in schools with fewer poor students.
Poverty and “Student Growth” in NJ

New Jersey Middle School SGPs & Low Income
Grades 6 to 8 Schools


www.newyorkprincipals.org / @scfeeney
“Student Growth” and Proficiency in NJ


www.newyorkprincipals.org / @scfeeney
Contradictory Information

- At Wheatley, 92% of our 8th Graders take the Algebra Regents exam. All but 1 passed (99.1% passing rate)
  - 73.5% of these students passed the Algebra Regents at the “College and Career Readiness” score of 80

YET...

- Only 39% of our 8th graders “passed” the easier 8th grade mathematics assessment

IN FACT...

- 59% of the students categorized as not meeting standards on 8th grade test subsequently earned a score of at least 80 on the Algebra Regents

www.newyorkprincipals.org / @scfeeney
A Letter to Parents about Testing

www.newyork principals.org

www.newyork principals.org / @scfeeney
What We Know

1. NYS Testing Has Increased Dramatically
2. The Tests were Too Long
3. Ambiguous Questions Appeared throughout the Exams
4. Children have Reacted Viscerally to the Tests
5. The Low Passing Rate was Predicted
6. The College Readiness Benchmark is Irresponsibly Inflated
7. State Measures are Contradictory
8. Students Labeled as Failures are Forced Out of Classes
9. The Achievement Gap is Widening
10. The Tests are Putting Financial Strains on Schools
11. The Tests are Threatening Other State Initiatives
What We Do Not Know

1. How these tests will help our students
2. How these tests will help improve student skills or understanding
3. The underlying cause of low test scores
4. What to expect next year
5. How much this is costing already-strained taxpayers
So if not through tests, what does accountability look like?
Montgomery County, MD

Peer Assistance and Review (PAR) program

- Master teachers mentor new and struggling teachers
- Performance standards based on National Board for Professional Teaching Standards
- If mentoring does not work, PAR panel (8 teachers, 8 principals) can vote to fire a teacher
- Collaborative process built upon support and trust
- Largest school system in Maryland; 16th largest in nation (over 144,000 students)
Cincinnati Public Schools

Teacher Evaluation System (TES)

- Based on Danielson framework
- Mid-career teachers trained as Peer Evaluators, a role they assume for three years before returning to classroom
- 5-year rotation of a comprehensive evaluation
- Research indicates a strong correlation between TES scores for teachers and difference in student achievement growth
- Evaluations associated with a significant increase in student scores in mathematics
Massachusetts RTTT Model

Educator Evaluation Model

- Applies to all educators, including teachers, principals and superintendents
- Four performance standards (similar to Danielson), with demonstration items that can be provided by teacher
- Four performance ratings: Unsatisfactory, Needs Improvement, Proficient, Exemplary. Bands, NOT points are assigned to teachers
- The length and characteristics of an individual’s Educator Plan varies in length and characteristics: Improvement Plan, Directed Plan (1-yr), Self-Directed plan (1-yr), Self-directed plan (2-yrs)
- Trends of multiple measures of student learning outcomes given to all educators as a lens through which to view the overall evaluation. This Impact on Student Learning can change the type of educator plan

www.newyork principals.org / @scfeeney
Final Thoughts
What if they are wrong?

- How much damage has been done?
- How much money has been spent?
- How much are we changing schools?
- How are we changing student perceptions of school?
Leadership

- After we released our APPR letter, Commissioner King dismissed our concerns, stating that “It’s not at all surprising [that the introduction of a new evaluation system] would produce anxiety.” (NYT, 11/27/11)

- Chancellor Tisch blamed parents for passing test stress on to their children (Brian Lehrer Show, 8/15/13)

- King blamed NYSUT for the stress related to students taking the tests (Times-Union, 6/5/13)

- In his “New York State Assessment Update” presentation, King stated that the increase in testing was all “optional” and the result of local district decisions

www.newyorkprincipals.org / @scfeeney
Not everything that counts can be counted, and not everything that can be counted counts. (Einstein)
What gets measured gets done.
What gets nurtured grows.
Research Notes


Rogers, Thomas (September 2013) Testimony to the Senate Education Committee Hearing at Suffolk Community College on 9/17/2013 http://www.nysenate.gov/files/Dr.%20Thomas%20Rogers.pdf.


Samtani, Hiten (23 May 2012) “More Parents Are Saying No to Pearson’s Field Tests”
