SAYING IT WITH DATA: NAEP AND NEW JERSEY’S STATE ASSESSMENTS

National Assessment of Educational Progress (NAEP)

NJDOE Summer Conference
One Vision, Many Voices: Academic Achievement for ALL Students

Atlantic City, NJ
August 8, 2006

Barbara Smey-Richman, EdD.
New Jersey NAEP Coordinator
Hi! I’m from the “NJDOE” and I’m here to help!!
NAEP Means What?

NAEP: National Assessment of Educational Progress

NAEP is also called the Nation’s Report Card

NAEP Rhymes with “Grape”
Comparing NAEP with the New Jersey tests is like comparing ....?
Metaphor: The Garden State

Goats and Nectarines: Garden State Farm -- Comparison of Profits (2003 and 2005)

Annual Profit per Sector

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goats</td>
<td>$74,000</td>
<td>$72,000</td>
</tr>
<tr>
<td>Nectarines</td>
<td>$37,000</td>
<td>$38,000</td>
</tr>
</tbody>
</table>
Surprise: Same Data/Different Story

NAEP and NJ GEPA Language Arts Literacy (LAL): Different Tests Yield Different Results (2003-2005)

<table>
<thead>
<tr>
<th>Grade 8</th>
<th>NJ GEPA LAL</th>
<th>NAEP Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>74% 2003</td>
<td>37% 2003</td>
</tr>
<tr>
<td></td>
<td>72% 2005</td>
<td>38% 2005</td>
</tr>
</tbody>
</table>

Percent At or Above Proficient
NAEP and State Tests

Student Results on NAEP and State Tests:
Reading Grade 8 (2003-2005)

<table>
<thead>
<tr>
<th></th>
<th>NJ LAL 2003</th>
<th>NJ LAL 2005</th>
<th>NAEP Reading 2003</th>
<th>NAEP Reading 2005</th>
<th>Missouri Reading 2003</th>
<th>Missouri Reading 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent At or Above Proficient</td>
<td>74</td>
<td>72</td>
<td>37</td>
<td>38</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>
Session Themes:

- NAEP Characteristics
- NAEP and NJ Assessments
- NJ Student Results
- Achievement Gaps
- Simpson’s Paradox
- Final Observations
- Q & A
What is NAEP?

NAEP was authorized by Congress in 1969 to create a reliable way of determining areas of strengths and weaknesses in the American school system.
What is NAEP?

**NAEP** was created to answer the question:

Are students in American schools learning what they should be learning?
What is NAEP?

- NAEP’s mission is to produce national and state-level results for student populations.

- NAEP does not provide individual student, school or district results.
What's Tested?

- **NAEP** tests a variety of subjects
- The primary assessments are:
  - Reading
  - Math
  - Writing
  - Science
- **NAEP** also funds “special studies”
<table>
<thead>
<tr>
<th>Year</th>
<th>National</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Reading, Mathematics</td>
<td>Reading, Mathematics</td>
</tr>
<tr>
<td>2004</td>
<td>Foreign Language, Long-term Trend</td>
<td><em>(Cancelled)</em></td>
</tr>
<tr>
<td>2005</td>
<td>Reading, Mathematics, Science</td>
<td>Reading, Mathematics, Science</td>
</tr>
<tr>
<td>2006</td>
<td>U.S. History, Civics, Economics</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Reading, Mathematics, Writing</td>
<td>Reading, Mathematics, Writing</td>
</tr>
<tr>
<td>2008</td>
<td>Arts, Long-term Trend</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>Reading, Mathematics, Science</td>
<td>Reading, Mathematics, Science</td>
</tr>
<tr>
<td>2010</td>
<td>World History, Geography</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>Reading, Mathematics, Writing</td>
<td>Reading, Mathematics, Writing</td>
</tr>
<tr>
<td>2012</td>
<td>Civics, Foreign Language, Long-term Trend</td>
<td></td>
</tr>
</tbody>
</table>
NAEP Has Many Components

The National Assessment of Educational Progress

MAIN

National
Public & Nonpublic
Grades 4, 8, & 12

State
(Public)
(Grades 4 & 8)

District - Trial
(Public)
(Grades 4 & 8)

LONG-TERM

National
Public & Nonpublic
9, 13, & 17 yr olds
What is State NAEP?

State NAEP (1990) is an estimate of student performance for the

- 50 states
- Washington, DC
- Puerto Rico
- Department of Defense Education Activities

State NAEP is a biennial assessment
What is State NAEP?

- With NCLB, Title I funding became contingent upon school participation in NAEP at grades 4 and 8.
- Prior to 2003, school participation was voluntary.
- The NAEP data are not reported if the school-participation rate is low.
Who’s Tested?

- NAEP uses a complicated Multi-Stage Stratified Random Sampling method
- Sample sizes
  - Approximately 3,000 students are selected for each subject and grade level (same for all states)
  - NAEP selects about 3% of NJ’s 100,000 students for each grade-level cohort
How are Students Assessed?

“Matrix Sampling”

- Each student answers only a small portion of the NAEP items
- Students in the same classroom receive different booklets
What Content is Tested?

- NAEP does **not** have curriculum content standards.
- NAEP’s *assessment frameworks* are analogous to the New Jersey *test specifications*.
- The NAEP assessment frameworks include:
  - theoretical basis of the test
  - directions for item development
NAEP’s Theoretical Basis *

- **Context for Reading**
  - Literary Experience
  - Information - retrieving information
  - Perform a Task - following instructions

- **Aspects of Reading**
  - General Understanding
  - Interpretations
  - Reader/Text Connection
  - Context and Structure

* See the NAEP web site for further details.
What’s New With NAEP?

- **NAEP 2009** -- a totally new NAEP Reading Assessment
- The current NAEP trend line will be broken in 2009
Check It Out!!

NAEP Releases Test Items!!

Obtain a copy of the Questions Tool Quick Reference at this session
Add to Your Favorites List!

On the NAEP web site

http://nationsreportcard.gov

- Find sample questions *(NAEP Question Tool)*
- Review scoring guides
- Compare state groups *(NAEP Data Explorer)*
- View results by scores and achievement levels

NJ NAEP Web pages:
http://www.nj.gov/njded/assessment/naep
NAEP and New Jersey Assessments
Getting Back to “Why” So Different?

NAEP and NJ GEPA Language Arts Literacy (LAL): Different Tests Yield Different Results (2003-2005)

Grade 8

Percent At or Above Proficient

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>NJ GEPA LAL</td>
<td>74%</td>
<td>72%</td>
</tr>
<tr>
<td>NAEP Reading</td>
<td>37%</td>
<td>38%</td>
</tr>
</tbody>
</table>
Let’s Take a Closer Look!

What are the similarities and differences?
NAEP’s Definition of Reading

Reading is defined as an active process that involves

➤ Understanding written text

➤ Developing and interpreting meaning, and

➤ Using meaning as appropriate to the type of text, purpose, and situation

(NAEP 2009 Framework)
What's Being Tested?

Similarities?

Don’t NAEP and the NJ tests measure a common core of reading ability?
This diagram is only an approximation of how the NAEP reading and New Jersey LAL tests relate to each other. The diagram is not drawn to "scale," and the extent to which the tests overlap is not based on an alignment study.
An Obvious Difference

Language Arts Literacy

- Combined Tests
- Combined Scores only

Reading and Writing

- Separate R/W Tests
- Separate R/W Scores
Different Schools of Thought??

- NAEP’s assessment frameworks are developed as a result of a national-based process and reflect a national view.

- New Jersey’s content standards and test specifications for NJASK 4 and GEPA reflect the New Jersey perspective.
An Important Question - -

Are the NJ state scores higher because NJ’s instruction is focused on the LAL standards?
NAEP/NJ Differences

- **Test length**
  - Content questions take 50 minutes per student; total test would take 7 hours

- **Content and skills assessed**
  - NAEP tests a wider domain of knowledge since the test is longer

- **Average scale score points for reading tests**
  - NAEP has a 500 pts. scale; NJ, 300 pts. scale

- **Student motivation**
  - NAEP has no consequences for the students
  - Are students focused and trying their best?
NAEP/NJ Differences

- **Standard-setting methods**
  - *NAEP’s* process for determining the achievement levels may be different from that used by the individual states
  - Cut points vary according to the method used

- **Item difficulty and format**
  - *NAEP* items may be more difficult
  - For NAEP, about half the time is spent on short/extended constructive response items
  - Also has multiple choice items
NAEP’s Achievement Levels

Level Descriptors

- **Basic**: Partial mastery of prerequisite knowledge/skills that are fundamental for proficient work *(At grade level)*

- **Proficient**: Solid academic performance -- demonstrated competency over challenging subject matter

- **Advanced**: Superior performance

[Below Basic: Incomplete knowledge/skills necessary for proficient work]
Play It Again Sam

1. NAEP and the NJ tests are inherently different tests
2. Test develops caution against making comparisons
3. Controversial Question: Can NAEP be used to verify overall trends found in state results?
What are the Implications?

Because of the inherent differences:

- Proficient performance is defined uniquely for each large scale assessment

- The definitions of “proficient” set by states and by NAEP have no observable agreement

* Linn, Robert, *Large-Scale Assessment Conference*, San Antonio, TX, June 2005
Over Interpreting the Data

By twelfth grade, Black students are typically four years behind white and Asian students. Black students are finishing high school with a junior high school education.*

What's an Achievement Gap?

**NAEP**  
Achievement Gap = A statistically significant difference in the scores between two groups

If not statistically significant, it is not a gap.

Johnny Cochran:  
If it doesn’t fit, you must acquit.
# What's an Achievement Gap?

Same difference in scale scores, but different decisions

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>SE</th>
<th>Group 2</th>
<th>SE</th>
<th>Point Difference Between Groups</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test 1 GAP</strong></td>
<td>210</td>
<td>2.5</td>
<td>200</td>
<td>3.0</td>
<td>10*</td>
<td>t = 2.56 p &lt; .05</td>
</tr>
<tr>
<td><strong>Test 2 Not a GAP</strong></td>
<td>210</td>
<td>3.5</td>
<td>200</td>
<td>4.5</td>
<td>10</td>
<td>t = 1.75 p &gt; .05</td>
</tr>
</tbody>
</table>

* Statistically significant
Pictures at An Exposition

Overall Results
We Have Reason To Cheer!!
How Do NJ 8\textsuperscript{th} Graders Compare?

<table>
<thead>
<tr>
<th>Reading Average Scale Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey's average reading scale score compared with scores for the nation and other participating jurisdictions, grade 8 public schools: 2005</td>
</tr>
</tbody>
</table>

NJ scale scores were:
- Lower than 1
- Equal to 16
- Higher than 34

NAEP Reading
How Do NJ 4th Graders Compare?

New Jersey's average reading scale score compared with scores for the nation and other participating jurisdictions, grade 4 public schools: 2006

NJ scale scores were:
- Lower than 3
- Equal to 24
- Higher than 24
How Do NJ 4<sup>th</sup> and 8<sup>th</sup> Graders Compare to the Nation?

At grades 4 and 8, NJ results were higher than the national public schools in 2003 and 2005.

* Statistically significant
Comparing Jurisdictions With Similar Demographics
Is Like States in Northeast?
How Do the Nation’s Census Regions Compare? Grade 4

NE Compared to Other Regions

NAEP 2003 & 2005

<table>
<thead>
<tr>
<th></th>
<th>MW</th>
<th>NE</th>
<th>S</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midwest</td>
<td>&lt;</td>
<td>&gt;</td>
<td>&gt;</td>
<td></td>
</tr>
<tr>
<td>(MW)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>&gt;</td>
<td>&gt;</td>
<td>&gt;</td>
<td></td>
</tr>
<tr>
<td>(NE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>&lt;</td>
<td>&lt;</td>
<td>&gt;</td>
<td></td>
</tr>
<tr>
<td>(S)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>&lt;</td>
<td>&lt;</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>(W)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How Do NJ 8th Graders Compare to State in the Northeast? (2005)

Note: New Jersey’s average scale (269) score was not different from the average scale score for the Northeast Census Region (267).
Are the Nation’s Reading Scores Improving? (Long-term Trend) (2004)

**Key Findings:**

**9-yr olds:** The 2004 score is higher than previous years.

**13-yr olds:** The scale score in 2004 is higher than in 1971 and in 1975, no different for subsequent years.

---

* Significantly different from 2004.

Did the NJ NAEP Reading Scores Improve from 2003 to 2005?

![Graph showing reading scores for Grade 4 from 1992 to 2005.](image)

**Grade 4**

No significant change in scales scores from 2003 to 2005.

**Grade 4**

<table>
<thead>
<tr>
<th>Category</th>
<th>Below Basic</th>
<th>At Basic</th>
<th>At Proficient</th>
<th>At Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>31</td>
<td>34</td>
<td>27</td>
<td>8</td>
</tr>
<tr>
<td>1992¹</td>
<td>35</td>
<td>32</td>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>1994¹</td>
<td>30</td>
<td>31</td>
<td>28</td>
<td>11</td>
</tr>
<tr>
<td>2003</td>
<td>32</td>
<td>31</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>2005</td>
<td>31</td>
<td>34</td>
<td>27</td>
<td>8</td>
</tr>
</tbody>
</table>

¹ Accommodations were not permitted for this assessment.

No significant change in the percentage At or Above Proficient from 2003 to 2005.
Did the NJ NAEP Reading Scores for Improve from 2003 to 2005?

No significant change in scales scores from 2003 to 2005.

No significant change in the percentage At or Above Proficient from 2003 to 2005.
What's the Trends for NJASK 4?

New Jersey Assessment of Fourth Grade Students
Percent Proficient and Above by Content Areas (1999-2005)

1Elementary School Proficiency Assessment (ESPA), 1999-2002; New Jersey Assessment of Skills and Knowledge (NJ ASK4), 2003-2005. For Language Arts Literacy, 2001 was the standard setting year for the ESPA and NJ ASK4. For Mathematics, 1999 was the standard setting year for the ESPA and NJ ASK4.
What's the Trends for GEPA?

New Jersey Grade Eight Proficiency Assessment
Percent Proficient and Above for All Students by Content Areas (1999-2005)
Pictures at an Exposition

Achievement for ALL Students

• Students with Disabilities (SD)
• English Language Learner (ELL)
NAEP SD/Regular Education for NJ, Grade 4, (2005)

Students with Disabilities (SD)

<table>
<thead>
<tr>
<th>Category</th>
<th>Below Basic</th>
<th>At Basic</th>
<th>At Proficient</th>
<th>At Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>70</td>
<td>23</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>27</td>
<td>32</td>
<td>30</td>
<td>11</td>
</tr>
</tbody>
</table>

NAEP SD/Regular Education for NJ & Other Jurisdictions, Grade 4 (2005)

New Jersey, Nation, and Northeast

<table>
<thead>
<tr>
<th>Category</th>
<th>Below Basic</th>
<th>At Basic</th>
<th>At Proficient</th>
<th>At Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Public</td>
<td>67</td>
<td>22 ✦</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>New Jersey</td>
<td>70</td>
<td>23 ✦</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Northeast (Census)</td>
<td>64</td>
<td>25 ✦</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>No Regular Ed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Public</td>
<td>34</td>
<td>34</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>New Jersey</td>
<td>27</td>
<td>32</td>
<td>30</td>
<td>11</td>
</tr>
<tr>
<td>Northeast (Census)</td>
<td>25</td>
<td>35</td>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>

* Statistically significant

At Basic
Northeast Region was higher than NJ and Nation
No other comparisons were significant

ELL for NJ Not Reported

- The NAEP ELL results for NJ were unreliable because:
  - Small sample sizes
  - High Standard Errors (SE)
- NAEP sample sizes for NAEP are small due to exclusionary rules used since 1992
- NAEP 2007 -- exclusionary rule will be updated and aligned with the NJ rules
Pictures at an Exposition

Achievement for ALL Students

Students from Low-Income Families

(Free/Reduced Lunch)
Low-Income Gap in NJ NAEP, At/Above Basic, Grade 4 (2003 and 2005)

NJ Eligible for Free or Reduced Lunch

<table>
<thead>
<tr>
<th>Category</th>
<th>Below Basic</th>
<th>At Basic</th>
<th>At Proficient</th>
<th>At Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 Eligible</td>
<td>54</td>
<td>31</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>2005 Eligible</td>
<td>55</td>
<td>28</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Not eligible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 Not eligible</td>
<td>20</td>
<td>32</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td>2005 Not eligible</td>
<td>22</td>
<td>32</td>
<td>33</td>
<td>12</td>
</tr>
</tbody>
</table>

Eligible < Non Eligible

For all Comparisons

For all Comparisons

<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott District 2004</td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td>190</td>
</tr>
<tr>
<td>All Districts</td>
<td>200</td>
</tr>
<tr>
<td>All Schools</td>
<td>204</td>
</tr>
<tr>
<td>Non-Abbott District 2004</td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td>190</td>
</tr>
<tr>
<td>All Districts</td>
<td>220</td>
</tr>
<tr>
<td>All Schools</td>
<td>221</td>
</tr>
</tbody>
</table>

From the Instructional Data Management System, [www.idms.com](http://www.idms.com)
NJASK 4 & the District Factor Groups

**NJASK Grade 4 LA 2004**
- Scale Score
- District Factor Group

<table>
<thead>
<tr>
<th>Category</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Factor Group A 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group B 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group CD 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group DE 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group FG 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group GH 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group I 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>District Factor Group J 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Districts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Low SES**

**High SES**

*Source: N.J Department of Education, NJASK, 2004 LA Assessments*

From the Instructional Data Management System, www.idms.com
Pictures at an Exposition

Achievement for ALL Students

Gender
Gender Achievement Gap

NAEP Grade 4: Average New Jersey Reading Scale Scores (1992, 1994 and 2003)

New Jersey

Girls > Boys (7 pts) in 2003

No Gap Change from 1992 to 2005

* Females significantly higher than males.  N – Accommodations not permitted in 1992 and 1994,
Gender Gap/Gap Change from 2003 to 2005, Grade 8

At or Above Proficient
2003 - 2005

➢ Females > Males

No change in gender gap from 2003 to 2005

From the Instructional Data Management System, www.idms.com

# Gender Gap: NJ, US, NE, Grade 8

## Scale Scores, Grade 8 Reading, 2005

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th>Males</th>
<th>Scale Score Difference</th>
<th>Male/Female Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey</td>
<td>273</td>
<td>266</td>
<td>-8</td>
<td>Significant</td>
</tr>
<tr>
<td>National Public</td>
<td>266</td>
<td>255</td>
<td>-11</td>
<td>Significant</td>
</tr>
<tr>
<td>Northeast Census Region</td>
<td>272</td>
<td>263</td>
<td>-9</td>
<td>Significant</td>
</tr>
</tbody>
</table>
Pictures at an Exposition

Achievement for ALL Students

Race/Ethnicity
Race Gap: NAEP Achievement Levels
Grade 8 (2003 and 2005)

Percentages of students at each achievement level for reading, grade 8
Race/ethnicity used in NAEP reports after 2001 [SDRACE]
New Jersey, 2003 and 2005

<table>
<thead>
<tr>
<th>Category</th>
<th>Below Basic</th>
<th>At Basic</th>
<th>At Proficient</th>
<th>At Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>White 2003</td>
<td>12</td>
<td>42</td>
<td>42</td>
<td>4</td>
</tr>
<tr>
<td>White 2005</td>
<td>12</td>
<td>40</td>
<td>43</td>
<td>5</td>
</tr>
<tr>
<td>Black 2003</td>
<td>42</td>
<td>43</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Black 2005</td>
<td>38</td>
<td>48</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic 2003</td>
<td>39</td>
<td>44</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic 2005</td>
<td>35</td>
<td>51</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Asian Amer/Pacif 2003</td>
<td>8</td>
<td>30</td>
<td>51</td>
<td>12</td>
</tr>
<tr>
<td>Asian Amer/Pacif 2005</td>
<td>5</td>
<td>28</td>
<td>54</td>
<td>12</td>
</tr>
</tbody>
</table>

NAEP 2005 --Significance
Percent At or Above Proficient

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>B</th>
<th>H</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>White (48%)</td>
<td>&gt;</td>
<td>&gt;</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>Black (15%)</td>
<td>&lt;</td>
<td>=</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>Hispanic (14%)</td>
<td>&lt;</td>
<td>=</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>Asian (66%)</td>
<td>&gt;</td>
<td>&gt;</td>
<td>&gt;</td>
<td></td>
</tr>
</tbody>
</table>
Race (1992-2005): Hispanics, Grade 4

Good News:
No change from 2003 to 2005.

* Statistically significant

Average scale scores for reading, grade 4
Raceneutral in NAEP reports after 2001 [SDRACE] = Black

NOTE: The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.
Race (1992-2005): Asian/PI, Grade 4

**Good News:**
Scores higher in 2005 than in 1992
No change from 2003 to 2005

* Statistically significant

NOTE: The NAEP Reading score ranges from 0 to 500. Observed differences are not necessarily statistically significant.
Race (1992-2005): White and Black Results, Grade 4

White Students – No Change

Black Students – No Change

NOTE: The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.
Race Gaps: NJ and Nation, Grade 4

Average Scale Scores
For Race/Ethnicity, 2005

<table>
<thead>
<tr>
<th>Category</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>190</td>
</tr>
<tr>
<td>New Jersey White</td>
<td></td>
</tr>
<tr>
<td>New Jersey Black</td>
<td></td>
</tr>
<tr>
<td>New Jersey Hispanic</td>
<td></td>
</tr>
<tr>
<td>New Jersey Asian/PI</td>
<td></td>
</tr>
<tr>
<td>Nation (Public) White</td>
<td></td>
</tr>
<tr>
<td>Nation (Public) Black</td>
<td></td>
</tr>
<tr>
<td>Nation (Public) Hispanic</td>
<td></td>
</tr>
<tr>
<td>Nation (Public) Asian/PI</td>
<td></td>
</tr>
</tbody>
</table>

Note: The NAEP reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.


* Statistically significant
How Do NJ’s Black Students Compare?

Grade 4, 2005
Only 6 states have higher achievement for Black students
Why Flat Trend Data?
Digging Further?

Sometimes digging behind the “averages” can uncover an entirely different story.

Simpson’s Paradox

- Simpson’s Paradox is when the aggregate group score shows one pattern, but the subgroup scores show a different pattern.
- This phenomenon is seen in NAEP and SAT scores over time.
- Subgroups show significant gains over time, but the overall national or state averages are flat.
Simpson’s Paradox: NJ’s Trend
Data for Race, Grade 4

<table>
<thead>
<tr>
<th></th>
<th>% of Testing Pop.</th>
<th>Av. Scale Score</th>
<th>Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>All NJ Students</td>
<td>1992 100%</td>
<td>223</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2005 100%</td>
<td>223</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1992 69%</td>
<td>233</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>2005 58%</td>
<td>232</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1992 16%</td>
<td>198</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>2005 17%</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1992 11%</td>
<td>195</td>
<td>+11</td>
</tr>
<tr>
<td></td>
<td>2005 16%</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1992 4%</td>
<td>231</td>
<td>+10</td>
</tr>
<tr>
<td></td>
<td>2005 8%</td>
<td>241</td>
<td></td>
</tr>
</tbody>
</table>

- Why no gains overall?
- **Whites** -- Proportion of high scoring whites decreased, pulling down the average
- **Hispanics** -- Scores jumped, but averages are relatively low (206)
- **Asians** -- Scored jumped, but still proportion of the total population is low
Simpson's Paradox

Another example of a Simpson's Paradox was provided by Howard Wainer


Examined performance in New Jersey in comparison to other states and the nation
Simpson’s Paradox: NJ Compared to VT and NH, Grade 4 (2005)

It appears that 4th grade students in NH, and VT do better in reading than their counterparts in New Jersey.
Simpson’s Paradox: NJ Compared to VT and NH, Grade 4 (2005)

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th>Black</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey</td>
<td>58%</td>
<td>17%</td>
<td>16%</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>94%</td>
<td>1%</td>
<td>2%</td>
</tr>
<tr>
<td>Vermont</td>
<td>96%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>

- It is caused by the differences in the proportional representation of the ethnic groups.
- Each state's mean score is a product of the mean score within each ethnic group and its proportional representation in the population.
When we examine the average scale scores by ethnic group we find

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th>Black</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey</td>
<td>232</td>
<td>199</td>
<td>206</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>228</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Vermont</td>
<td>227</td>
<td>++</td>
<td>++</td>
</tr>
</tbody>
</table>

++ Reporting standards not met
Simpson’s Paradox

- Overall national and state results are informative at a general level
- The proportional representation of subpopulations should be considered
- To truly understand what students are learning, one must examine how the subgroups are performing
What Have We Learned?

This session began and ended with analyses of state-level results.
What Have We Learned?

NAEP’s original question:

Are students in American schools learning what they should be learning?

What do you think?
What Can NAEP Tell Us?

What did we learn about

- NJ’s overall reading performance?
- NJ’s subgroup reading performance?
  - Students with Disabilities
  - Students from Low-Income Families
  - Males and Females
  - White, Black, Hispanic and Asian/Pacific Islander students
Limitations to the NAEP Data

NAEP Data cannot tell us:

- Why gaps exist
- Policies for closing gaps
- Suggestions for interventions to improve student learning
Where to Look for Answers - Educational Research, for example:

- Exemplary schools -- successful with low-performing students
- What Works - a knowledge-base on effective schools and successful practices in the classroom
CONTACT INFORMATION


Barbara Smey-Richman, EdD
New Jersey Department of Education
Office of Evaluation and Assessment
P.O. Box 500
Trenton, NJ 08625-0500

Tel: (609) 984-1540
Fax: (609) 984-6032
Email: Barbara.Smey-Richman@doe.state.nj.us
Time to Hit the Beach?

Thanks for Your Interest in NAEP!!