Fourth- and Eighth-Grade NAEP: Mathematics Trends in the 21st Century

Discussant Comments

Glen Blume
Questions

• What questions arise from this work?
  – Questions I have
  – Questions you might wish to raise
  – “Next-steps” questions that might be investigated
Can we trust the results?

• Items
  – Algebra: 48 items, many released, cover many topics
  – Released items are representative; access to secure items offers “excellent sense of student performance”

• Research team
  – Thorough, experienced, able to obtain access

• Items
  – “Only” 48 items (5 per topic tests only 10 algebra topics)
  – They’re “released” items; authors try to say something meaningful without divulging item content

• Research team
  – Design, develop, collect, analyze, interpret
Can we trust the results?

• Significance
  – 2% $\rightarrow$ significant ($p < .01$) and
  – 4% $\rightarrow$ significant ($p < .001$)

• Item format
  – MC items $\rightarrow$ more items
    per unit of time

• Significance
  – Multiple tests raise the likelihood of accepting spurious results

• Item format
  – Most items are MC
    • MC test-taking strategies
    • Use of content other than that being tested to answer the question
Are these results important?

- Are statistically significant increases of 2% or more educationally significant?
  - If 26 per 100 students understood algebra last year and 28 per 100 this year, are we happy?
- Students can do but can’t explain. What does this predict for pursuing postsecondary mathematics and performing in the workforce?
- Students can do simple items but not complex ones.
Are these results important?

- Over the past 20 years, 4\textsuperscript{th} and 8\textsuperscript{th} gains translate to 2 grade levels!
- Individual item performance increases vs. declines point to a strong case for progress being made.
Are these results important?

Number of Algebra items with performance increases across time and number with declines across time

<table>
<thead>
<tr>
<th>Table</th>
<th># items</th>
<th># increases ≥ 2%</th>
<th># declines ≥ 2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>17</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Table 2</td>
<td>20</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Table 3</td>
<td>13</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Table 4</td>
<td>33</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Table 5</td>
<td>6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Table 6</td>
<td>13</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Table 7</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Table 8</td>
<td>23</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>
Why are these results happening?

- Even though content is not taught between Gr. 4 and Gr. 8, performance increases. Why?
- Why do we see greater Algebra gains at Gr. 8 than at Gr. 4? (Does it come only from the push to complete algebra in Gr. 8?)
- Is low performance due to lack of content knowledge or due to lack of computational fluency?
- Algebra: Find the equation of a line. Increase 55% → 65%. Why? Is technology use a factor in the increase?
- Item content, format, context, wording, symbolism affect performance. Which influences performance most? How does one balance these?
What are the implications for policy?

• In the Common Core era, ...
  – In what ways should NAEP evolve?
  – How can the competing goals—longitudinal data & current classroom emphases—be balanced?
  – Difficulty in mapping NAEP algebra items to CCSS
  – “Data from those [NAEP/CCSS matching items] suggest that it will be difficult to meet those standards in the short term” (p. 12).
  – Will NAEP evolve to match the CCSS Standards for Mathematical Practice? What do we know about how these match? How can we measure aspects of the SMP?
What are the implications for policy?

- Curricular emphasis
  - Why is there a Gr. 8 drop in Geom/Meas emphasis?
  - What, if anything, should be done in response?
  - From where has the Gr. 4 additional emphasis on Geom/Meas come? From Number/Operations? Is there an effect from technology use?
  - Data/Stat/Prob portion asks “Will NAEP evolve to match CCSSM emphasis and content organization?”
What are the implications for *policy*?

- “... ability to construct arguments, interpret a model, and explain reasoning are generally poor at both grades 4 and 8” (p. 12).
- Will *n* years of CCSS compliance remedy that? If so, how large is *n*?
- Elephant in the room: Achievement gaps
Miscellaneous questions

• What is *algebraic reasoning* according to:
  – The authors? CCSS-M? You?
  – Evidencing the CCSS-M SMPs in the context of algebra?

• Terms used:
  – Title: *Algebraic Reasoning*
  – Abstract: *Algebra knowledge*
  – p. 12: ... “mapping items to CCSS *algebra skills*”
  – P. 13, concluding section: *Algebraic thinking skills*
Miscellaneous questions

• (Alg. p. 12) “One could argue that many students have never been expected to explain their thinking.”
  – Is that true?

• (Alg. p. 13) Students need to be aware of novel representations (e.g., beam balance) for high-stakes assessments.
  – Maybe not, if what is being tested is application of curricular content to novel situations to test deeper understanding
Next-steps questions

• What else should researchers examine with respect to performance on NAEP?
  – Who will do it?
  – Who will fund it?
  – Will additional information inform curriculum, teaching, assessment, policy, and learning?