Closing the Engagement Gap: 
An Analysis of Student Effort and Engagement on Large-Scale Performance Assessments

Amy L. Abbott, Ph.D. 
Old Dominion University 

Douglas G. Wren, Ed.D. 
Old Dominion University 

Doris Zahner, Ph.D. 
Council for Aid to Education
Agenda

• Rationale
• College and Work Readiness Assessment (CWRA+)
• Integrated Performance Task (IPT)
• Student engagement on the IPT
• Effort and engagement on the CWRA+
• College to career
Rationale

• Research on students’ preferences for assessments has focused primarily on students at the undergraduate level (e.g., Birenbaum, 1997; Gijbels & Dochy, 2006; van de Watering, Gijbels, Dochy, & van der Rijt, 2008).

• “Classroom evaluation affects students in many different ways. For instance, it ... affects their motivation” (Crooks, 1988, p. 467).

• Assessment research since Crooks has failed to focus on younger students’ perceptions of assessment and how these perceptions affect their learning and motivation (MacMillan & Turner, 2014).
CWRA+
CWRA+ Performance Task

- **Real-world scenario** – A decision based on available finances

- **Specific Role** - Advisor to the City Manager

- **Decision to be made** - Which program should the city continue to fund?

- **Authentic student-produce work** – Written recommendation

- **Stakes** - A recommendation to the City Manager must be provided for the town meeting

- **Opposition** - Both are popular programs for the schools/students
If Tiverton cannot afford to fund the Forest Adventures and the Sports & School Experience programs at their current levels, what should the city do? Keep one program? Modify the program(s)? Something new?
If Tiverton cannot afford to fund the Forest Adventures and the Sports & School Experience programs at their current levels, what should the city do? Keep one program? Modify the program(s)? Something new?
If Tiverton cannot afford to fund the Forest Adventures and the Sports & School Experience programs at their current levels, what should the city do? Keep one program? Modify the program(s)? Something new?
If Tiverton cannot afford to fund the Forest Adventures and the Sports & School Experience programs at their current levels, what should the city do? Keep one program? Modify the program(s)? Something new?
If Tiverton cannot afford to fund the Forest Adventures and the Sports & School Experience programs at their current levels, what should the city do? Keep one program? Modify the program(s)? Something new?”
CWRA+ Scoring Rubric

Student responses are scored at one of six levels for each of the following skills:

• Analysis and Problem Solving
• Writing Effectiveness
• Writing Mechanics
Selected-Response Questions

25 Selected-Response Questions (SRQs)

- Scientific & Quantitative Reasoning
- Critical Reading & Evaluation
- Critique an Argument
INTEGRATED PERFORMANCE TASK
Integrated Performance Task

• *Compass to 2015*: “Our primary focus is on teaching and assessing those skills our students need to thrive as 21st century learners, workers, and citizens.”

• The *Integrated Performance Task* (IPT) is a series of elementary and middle school performance tasks designed to measure specific 21st century skills. The IPT is modeled after *College and Work Readiness Assessment* (CWRA+) performance tasks.

• Like CWRA+ performance tasks, each IPT involves a real-life problem and supplemental documents for students to evaluate before composing a solution to the problem.
Grade 4 IPT Scenario

You are a 4th-grade student at Smith Elementary School. Your principal, Mr. Beach, received two letters from concerned adults about Smith Elementary School’s nickname.

One letter is from some parents who do not like the Smith Elementary School nickname. These parents think the nickname should be changed. The other letter is from a grandparent who likes the nickname. She does not want it to be changed.

Mr. Beach wants the students to tell him which nickname to use. You should read and think about the information in this booklet. Then choose the nickname that would be best for Smith Elementary School.
Grade 7 IPT Scenario

Middleton Public Schools (MPS) has been looking at a proposal to change the starting and ending times for middle schools and high schools next year. Right now middle school hours are 9:30 a.m. to 4:00 p.m. The current high school hours are 7:30 a.m. to 2:00 p.m.

The proposed plan would switch the schedules so middle school would begin at 7:30 a.m. and end at 2:00 p.m. and high school would start at 9:30 a.m. and end at 4:00 p.m. If the plan is rejected, then the school hours will stay the same.

You work as an assistant to the superintendent of MPS. The superintendent is in charge of every school in Middleton. She wants you to read four documents and make a recommendation to her.

This Booklet Contains:

<table>
<thead>
<tr>
<th>Document</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document 1 - MPS School Hours Committee Report</td>
<td>2</td>
</tr>
<tr>
<td>Document 2 - Research Article: How Much Sleep Do Teens Need?</td>
<td>3</td>
</tr>
<tr>
<td>Document 3 - Social Networking Site Page/Online Discussion Board</td>
<td>4</td>
</tr>
<tr>
<td>Document 4 - Online News Story: Middleton Public Schools Will Switch School Hours</td>
<td>6</td>
</tr>
<tr>
<td>Prompt 1</td>
<td>7</td>
</tr>
<tr>
<td>Prompt 2</td>
<td>9</td>
</tr>
<tr>
<td>Glossary</td>
<td>11</td>
</tr>
<tr>
<td>IPT Rubric</td>
<td>13</td>
</tr>
</tbody>
</table>
IPT Scoring Rubric

Student responses are scored at one of four levels for each of the following skills:

• Critical Thinking

• Problem Solving

• Written Communication
IPT RESULTS
Survey Question

After completing their responses to the two IPT prompts, all fourth- and seventh-grade students were asked an additional question:

• Do you like taking this kind of test better than taking multiple-choice tests? Answer YES or NO. (You do not need to explain your answer.)
28.41% of 3,414 fourth-grade students who answered “yes” gave at least one reason related to the theme of engagement.
Do you like taking this kind of test better than taking multiple-choice tests?

All grade 7 responses

\[ N = 5,063 \]

Yes responses

\[ n = 2,372 \quad (46.85\% \text{ of } N) \]

Yes with explanation

\[ n = 736 \quad (14.54\% \text{ of } N) \]

Engagement reason(s) given

\[ n = 616 \quad (12.17\% \text{ of } N) \]

25.97% of 2,372 seventh-grade students who answered “yes” gave at least one reason related to the theme of engagement.
Engagement Codes

• Fun
• Opinion/Explain
• Challenge
• Real-Life Situation
• Typing/Computer
• Writing
Engagement Codes

• **Fun** – An aspect of the IPT or the IPT itself was fun
  “yes I do like it was fun looking back in the story to find my answer.” (4th)
  “Yes its better you are more in to it and the situation also is fun.” (7th)

• **Opinion/Explain** – Opportunity to give an opinion or explain answer
  “YES! I just love the kind of test where you get to express your opinion!” (4th)
  “yes because you can explain it in your own words.” (7th)
Engagement Codes

• **Challenge** – The task was challenging
  
  “yes I really did cause it was a hard challenge.” (4th)
  
  “I personally like doing this form of test. It seems to boost my thinking skills.” (7th)

• **Real-Life Situation** – The IPT posed a realistic scenario
  
  “I love the real life problem.” (4th)
  
  “Yes, because it is much more real than fake. I like tests with real life situations instead of fake tests.” (7th)
Engagement Codes

- **Typing/Computer** – Chance to type or use computer
  
  “yes I do, because I get to type on the computer. I hope I have more test like it!” (4\textsuperscript{th})
  
  “yes because i love typing it is my addiction that is why i'm explaining my answer anyway.” (7\textsuperscript{th})

- **Writing** – Opportunity to write
  
  “I like this kind of test because I like having to write a letter to a audience.” (4\textsuperscript{th})
  
  “Yes I love writing because i can express my ideas in a creative manner.” (7\textsuperscript{th})
## Results – Overall

<table>
<thead>
<tr>
<th>Code</th>
<th>Grade 4 ((n = 970))</th>
<th>Grade 7 ((n = 616))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opinion/Explain</td>
<td>53.40%</td>
<td>72.89%</td>
</tr>
<tr>
<td>Typing/Computer</td>
<td>28.35%</td>
<td>11.20%</td>
</tr>
<tr>
<td>Fun</td>
<td>17.22%</td>
<td>5.52%</td>
</tr>
<tr>
<td>Challenge</td>
<td>15.15%</td>
<td>18.51%</td>
</tr>
<tr>
<td>Writing</td>
<td>10.00%</td>
<td>5.03%</td>
</tr>
<tr>
<td>Real-Life Situation</td>
<td>3.30%</td>
<td>7.63%</td>
</tr>
</tbody>
</table>
## Results by Gender

<table>
<thead>
<tr>
<th></th>
<th>Grade 4</th>
<th></th>
<th>Grade 7</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td></td>
<td>(n = 404)</td>
<td>(n = 566)</td>
<td>(n = 285)</td>
<td>(n = 331)</td>
</tr>
<tr>
<td>Opinion/Explain</td>
<td>49.01%</td>
<td>56.54%</td>
<td>66.32%</td>
<td>78.55%</td>
</tr>
<tr>
<td>Typing/Computer</td>
<td>25.00%</td>
<td>30.74%</td>
<td>10.88%</td>
<td>11.48%</td>
</tr>
<tr>
<td>Fun</td>
<td>21.53%</td>
<td>14.31%</td>
<td>5.61%</td>
<td>5.44%</td>
</tr>
<tr>
<td>Challenge</td>
<td>18.07%</td>
<td>13.07%</td>
<td>21.75%</td>
<td>15.71%</td>
</tr>
<tr>
<td>Writing</td>
<td>9.16%</td>
<td>10.60%</td>
<td>3.86%</td>
<td>6.04%</td>
</tr>
<tr>
<td>Real-Life Situation</td>
<td>3.47%</td>
<td>3.18%</td>
<td>6.67%</td>
<td>8.46%</td>
</tr>
</tbody>
</table>
# Results by Race – Grade 4

<table>
<thead>
<tr>
<th></th>
<th>Asian $(n = 69)$</th>
<th>Black $(n = 236)$</th>
<th>White $(n = 570)$</th>
<th>2 or more $(n = 86)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opinion/Explain</td>
<td>59.42%</td>
<td>53.39%</td>
<td>53.33%</td>
<td>56.98%</td>
</tr>
<tr>
<td>Typing/Computer</td>
<td>28.99%</td>
<td>33.90%</td>
<td>26.67%</td>
<td>26.74%</td>
</tr>
<tr>
<td>Fun</td>
<td>13.04%</td>
<td>16.95%</td>
<td>18.25%</td>
<td>16.28%</td>
</tr>
<tr>
<td>Challenge</td>
<td>24.64%</td>
<td>12.29%</td>
<td>14.91%</td>
<td>15.12%</td>
</tr>
<tr>
<td>Writing</td>
<td>5.80%</td>
<td>11.02%</td>
<td>11.05%</td>
<td>8.14%</td>
</tr>
<tr>
<td>Real-Life Situation</td>
<td>4.35%</td>
<td>0.85%</td>
<td>4.04%</td>
<td>3.49%</td>
</tr>
</tbody>
</table>
## Results by Race – Grade 7

| Grade 7      | Asian  
|--------------|--------|
|              | (n = 31) | (n = 616) | (n = 165) | (n = 339) | 2 or more  
|              | (n = 72) |
| Opinion/Explain | 70.97%  | 75.76%  | 69.32%  | 79.17%  |  
| Typing/Computer | 9.68%   | 13.94%  | 10.91%  | 6.94%   |  
| Fun          | 0.00%   | 7.88%   | 5.60%   | 2.78%   |  
| Challenge    | 9.68%   | 13.33%  | 22.12%  | 18.06%  |  
| Writing      | 0.00%   | 4.85%   | 5.60%   | 5.56%   |  
| Real-Life Situation | 6.45%   | 7.27%   | 7.96%   | 6.94%   |
Conclusions

• Fourth- and seventh-grade students are inclined to report that the most engaging component of the IPT is being able to state their opinion and/or explain their answers on open-ended prompts.

• Fourth-grade students tend to report that the next most engaging IPT component is typing or having the chance to work on a computer.

• Seventh-grade students have a tendency to report that the next most engaging component of the IPT is the challenge presented by the assessment.
Conclusions – Gender

• Female students in grades 4 and 7 are more likely than male students to report that explaining their answers or giving their opinion is an engaging part of the IPT.

• Male students in grades 4 and 7 are more likely than female students to report that the challenge presented by the IPT is engaging.

• Fourth-grade male students are also more likely than female students to mention that the IPT is fun.

• Fourth-grade female students are more likely than male students to mention that typing or working on the computer while taking the IPT is engaging.
Conclusions – Race

• African-American students at both grade levels are more likely than their counterparts in the other racial groups to report that typing or working on a computer was an engaging component of the IPT.

• Asian students in grade 4 are more likely than their counterparts in the other racial groups to report that the challenge of the IPT was engaging for them.

• White students in grade 7 are more likely than their counterparts in the other racial groups to report that the challenge of the IPT was engaging for them.
CWRA+ RESULTS
Participants

16,441 students

• M = 16.70 years; SD = 1.31 years

• 48% (7,943) Male; 48% (7,963) Female; 2% Decline to State

• Race
  – 57% (9,375) White, non-Hispanic
  – 15.4% (2,532) African-American/Black, non-Hispanic
  – 8% (1,375) Asian
  – 8% (1,349) Hispanic or Latino
  – 11.6% Other or Decline to State – not used in analysis

• 89.9% (14,797) English Primary Home Language
Schools

166 institutions

• 50.7 % (84) Public
• 49.3% (82) Independent/Private
Survey

1. How much effort did you put into the written-response task?
2. How much effort did you put into the selected-response questions?
3. How engaging did you find the written-response task?
4. How engaging did you find the selected-response questions?
Survey

1 = No effort at all
2 = A little effort
3 = A moderate amount of effort
4 = A lot of effort
5 = My best effort
Engagement PT

![Bar Chart](chart.png)
Engagement SRQs

![Bar chart showing the count of Engagement SRQs for different values of ENGAGING_SR. The chart displays bars for values 1 to 5 with the count ranging from 0 to 6,000.]
## Results

<table>
<thead>
<tr>
<th></th>
<th>Effort**</th>
<th>Engaging**</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT</td>
<td>3.79 (.99)</td>
<td>2.93 (1.06)</td>
</tr>
<tr>
<td>SRQ</td>
<td>3.24 (1.04)</td>
<td>2.40 (1.07)</td>
</tr>
</tbody>
</table>
RESULTS BY GENDER
Engagement by Gender

**ENGAGING_PT**

- **SEX**
  - Male
  - Female

**ENGAGING_SR**
## Results – Survey Questions

<table>
<thead>
<tr>
<th></th>
<th>Effort PT**</th>
<th>Effort SRQ**</th>
<th>Engage PT</th>
<th>Engage SRQ**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>3.86 (1.00)</td>
<td>3.16 (1.06)</td>
<td>2.94 (1.09)</td>
<td>2.42 (1.09)</td>
</tr>
<tr>
<td>Females</td>
<td>3.90 (.96)</td>
<td>3.32 (1.01)</td>
<td>2.93 (1.04)</td>
<td>2.37 (1.05)</td>
</tr>
</tbody>
</table>
## Results – CWRA+ Scores

<table>
<thead>
<tr>
<th></th>
<th>PT Score**</th>
<th>SRQ Score**</th>
<th>Total Score**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td>984 (204)</td>
<td>1018 (214)</td>
<td>1013 (183)</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td>1029 (199)</td>
<td>1030 (198)</td>
<td>1037 (174)</td>
</tr>
</tbody>
</table>
RESULTS BY RACE/ETHNICITY
Effort PT by Race/Ethnicity**

![Effort PT by Race/Ethnicity Graph](image)
Effort SRQs by Race/Ethnicity**
Engagement PT by Race/Ethnicity**
Engagement SRQ by Race/Ethnicity **
## Results – Survey Questions

<table>
<thead>
<tr>
<th></th>
<th>Effort PT**</th>
<th>Effort SRQ**</th>
<th>Engage PT**</th>
<th>Engage SRQ**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>3.69 (.97)</td>
<td>3.13 (.97)</td>
<td>3.03 (1.02)</td>
<td>2.46 (1.00)</td>
</tr>
<tr>
<td>African-American</td>
<td>3.71 (1.05)</td>
<td>3.27 (1.07)</td>
<td>2.86 (1.11)</td>
<td>2.53 (1.11)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3.72 (1.00)</td>
<td>3.14 (1.05)</td>
<td>2.89 (1.06)</td>
<td>2.35 (1.05)</td>
</tr>
<tr>
<td>White</td>
<td>3.85 (0.96)</td>
<td>3.28 (1.03)</td>
<td>2.96 (1.05)</td>
<td>2.37 (1.07)</td>
</tr>
</tbody>
</table>
## Results – CWRA+ Scores

<table>
<thead>
<tr>
<th></th>
<th>PT Score**</th>
<th>SRQ Score**</th>
<th>Total Score**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>1075 (187)</td>
<td>1068 (198)</td>
<td>1076 (165)</td>
</tr>
<tr>
<td>African-American</td>
<td>897 (184)</td>
<td>925 (184)</td>
<td>924 (161)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>979 (191)</td>
<td>985 (198)</td>
<td>992 (170)</td>
</tr>
<tr>
<td>White</td>
<td>1032 (200)</td>
<td>1055 (204)</td>
<td>1051 (175)</td>
</tr>
<tr>
<td>All</td>
<td>1010 (202)</td>
<td>1027 (206)</td>
<td>1028 (178)</td>
</tr>
</tbody>
</table>
## Results – Correlations

<table>
<thead>
<tr>
<th></th>
<th>Score PT</th>
<th>Score SRQ</th>
<th>Score Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort PT</td>
<td>.29**</td>
<td></td>
<td>.25**</td>
</tr>
<tr>
<td>Effort SRQ</td>
<td></td>
<td>.21**</td>
<td>.17**</td>
</tr>
<tr>
<td>Engage PT</td>
<td>.20**</td>
<td></td>
<td>.17**</td>
</tr>
<tr>
<td>Engage SRQ</td>
<td></td>
<td>.05**</td>
<td>.02</td>
</tr>
</tbody>
</table>
Conclusions

- Students tend to put more effort into the PT than the SRQs.
- Students tend to find the PT to be more engaging than the SRQs.
- Correlations between effort/engagement and CWRA+ performance were weak to moderate, although statistically significant.
Conclusions - Gender

- Females put more effort into the assessment than males.
- Males found the SRQs to be more engaging than females.
- Females had significantly higher CWRA+ scores than males.
- Correlations were similar to whole group results.
Conclusions – Race/Ethnicity

• Results were mixed with respect to race:
  – White students reported they put forth more effort.
  – Asian students reported the PT to be more engaging whereas African American students were more engaged with the SQRs.

• While Asian students put forth the least amount of effort, they were significantly more engaged and also had the highest CWRA+ scores.
Discussion

• Although students, overall, reported moderate to high levels of engagement and effort, it did not correlate strongly with test scores.

• Student ability is another factor that needs to be considered.
COLLEGE-TO-CAREER
Performance on the CLA+, a measure of critical thinking and written communication skills – skills that have been deemed as important in the workplace (Hart Research Associates, 2006 & 2013), is a significant predictor of being employed. Arum et al. (2012) found similar results.

There are actually some high-performing students from less selective or non-selective who need to have an opportunity to showcase their skills. And these students are often overlooked by employers because they do not come from the top or mid-level schools.
Participants

• College seniors – took CLA+ in spring 2014

• Surveyed 3, 6, and 12 months following graduation

• Approximately 13,000+ students
  – 1,585 agreed to participate in the survey
  – 993 persisted through all three phases
## Results – Post-College Outcomes

Predictive Validity of CLA+ and Other Variables on Post-college Outcomes

<table>
<thead>
<tr>
<th>CLA+ and</th>
<th>Post-college outcomes (all)</th>
<th>Salary</th>
<th>Employ</th>
<th>Full-time employ</th>
<th>Grad school</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLA+ only</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>EAA</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barron’s</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Field of study</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
Discussion

• CLA+ is predictive of post-college outcomes such as employment, salary, and enrollment in continuing education.

• Despite approximately 1.8 million individuals graduating each year (Hussar & Bailey, 2014), employers are still finding a skills gap (Hart Research Associates, 2015).

• Use of CLA+ scores can address some of the challenges employers and recent graduates face.

• Traditional career services and job-search resources typically do not provide students with a platform to demonstrate critical thinking skills to employers.
Implications

• Increase the use of challenging performance tasks with realistic scenarios and decrease the reliance on traditional assessments.

• Continue to increase the number of online assessments in which students construct responses to open-ended prompts.

• “What we choose to assess is what will end up being the focus of classroom instruction” (Gordon Commission, 2013, p. 9).

• The Every Student Succeeds Act of 2015 calls for “multiple measures of student achievement, including measures that assess higher-order thinking skills and understanding” (National Conference of State Legislatures, 2015, p. 3).
Further Research

- Determine the value that K-12 students place on different assessment types (e.g., traditional, performance-based).
- Analyze the relationship between students’ proficiency on elementary and middle school performance assessments and measures of high school success (e.g., GPA, on-time graduation).
- Examine the predictive validity of performance-based assessment results for the high school-to-college space.
- In the college-to-career space, collect data from employers to validate higher-order skills in the workplace.
References


Contact Information

Doris Zahner, Ph.D.
Director of Test Development & Measurement Science
Council for Aid to Education
DZahner@cae.org

Douglas G. Wren, Ed.D.
Assistant Adjunct Professor, Educational Foundations & Leadership
Old Dominion University
DWren@odu.edu

Amy L. Abbott, Ph.D.
Instructor, Department of Teaching & Learning
Old Dominion University
ALAbbott@odu.edu