Superscoring: Is it Fair?
During this session, you will learn more about:

1. Findings from ACT’s Higher Education Score Use Survey
2. The validity and fairness of superscoring for predicting college success
3. Impact of superscoring on subgroup differences
Presenters

Kenton Pauls  
*National Director, Higher Education at ACT*

Krista Mattern  
*Senior Director of ACT’s Validity and Efficacy Research Department*
Survey Asked

1. Policy when students provide multiple ACT or SAT test results
   • Superscore, highest composite, most recent, average

2. Policy for when students submit both SAT and ACT scores
   • Use concordance, not use concordance, take the most recent test (either one)

3. Rationale for current policy

4. Factors impacting superscoring policy adoption
   • Technical, other?

5. Campus receptivity/climate
Survey Revealed

• Approaches vary greatly

• Old ACT advice still driving current policy
  • Old ACT advice didn’t support superscoring so “…we never changed.”
  • Receptive to change in policy, but it will take time

2019 ACT Score Use Survey

Multiple ACT's
- Superscore: 34%
- Highest Score: 50%

Multiple SAT's
- Superscore: 49%
- Highest Score: 34%

15% points
Survey Revealed

- Most (but not all) use the latest concordance
  - 24% look at each test separately – no concordance

- Split/indifferent about ACT calculating superscore for institutions
  - Interest high for single report with full ACT test history

- Superscoring math
  - SAT sum easier than ACT Average
Other Notable Findings

- NAIA & NCAA different superscoring approaches
- Desire for disaggregated data remains
- Concerns about perception of privilege exist
- Openness to superscoring policy change
- Faculty play a role for this type of policy change
- Net tuition revenue risks
  - higher scores, more aid, less tuition
- Superscoring is a settled question…for some
TESTING, TESTING:
WHAT IS THE FAIREST SCORE WHEN APPLICANTS RETAKE ADMISSIONS TESTS?
Impetus

- Frequency of retesting is increasing
- Not consistent treatment of multiple scores across institutions
- Questions about fairness and equitable practices
- Dated literature; mainly focused on SAT
Previous Research

Differential Validity
- Slightly higher correlation coefficients for average score; $rs$ are .01 to .02 higher

Differential Prediction
- FYGPA for retesters was underpredicted
- Underprediction ↑ as testing occasions ↑
- Average method > underprediction; superscoring < underprediction
Research Questions

1. Which composite scoring method (average, highest, last, and superscoring) is most predictive of first-year grade point average?

2. Which composite scoring methods exhibits the least amount of differential prediction by number of testing occasions?
Current Study
Extends previous findings in four substantive ways:

1. Majority of research on SAT which is comprised of 2 scores; ACT has 4 subject tests
2. Previous research assumed non-repeaters were accurately predicted
3. Admission decisions are rarely based solely on test scores; ran models with HSGPA
4. Examined diversity implications for various scoring methods
Measures

**ACT tests scores.** English, mathematics, reading, and science – from all testing administrations were obtained from the student’s official ACT record.

1. *Last ACT Composite score.* This composite score reflects the score that the student earned on the last, or most recent, time they took the ACT.

2. *Average ACT Composite score.* This composite score is the average of all ACT Composite scores earned across test administrations/attempts, rounded to the nearest whole number.

3. *Highest ACT Composite score.* This composite score represents the highest ACT Composite score earned during a single administration.

4. *Superscored ACT Composite score.* This composite score takes the highest ACT subject test score (English, reading, mathematics, and science) across administrations and then computes the ACT Composite score for those highest subject test scores.

**Number of ACT Administrations.** Count of the number of times a student took the ACT during their sophomore through senior year of high school (M = 2.3): 1 time (29.1%), 2 times (35.3%), 3 times (20.2%), and 4 or more times (15.4%).

**High School Grade Point Average (HSGPA)** self-reported on ACT registration form; coursework taken and grades earned in English, math, social studies, and science (M = 3.40, SD = 0.50).

**First-Year Grade Point Average (FYGPA)** provided by participating colleges and universities (M = 2.73, SD = 0.95).
Table 1. Means, standard deviations, and intercorrelations of study variables

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Last</td>
<td>22.6</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Average</td>
<td>22.2</td>
<td>4.1</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Highest</td>
<td>22.9</td>
<td>4.2</td>
<td>0.98</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Superscored</td>
<td>23.3</td>
<td>4.2</td>
<td>0.97</td>
<td>0.97</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>HSGPA</td>
<td>3.40</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.51</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>FYGPA</td>
<td>2.73</td>
<td>0.95</td>
<td>0.40</td>
<td>0.39</td>
<td>0.40</td>
<td>0.41</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Note. N= 277,551. All correlations are significant at \( p < .0001 \). FYGPA = first-year grade point average.
Differential Prediction of Composite Scoring Method by Number of Testing Occasions
Differential Prediction of Composite Scoring Method by Number of Testing Occasions
Reduction of Differential Prediction with the Inclusion of HSGPA
Diversity implications of different scoring methods by three levels of selectivity (top 15%, 50%, and 85%).

Gender, ethnic, and income makeup is unaffected by the choice of scoring method.

- For top 15% selectivity scenario, all scoring methods resulted in an admitted class:
  - 45% male
  - 4% minority
  - 9-10% low-income.
Results can inform the college admission practice and policies

- Validity
  - Superscores showed the strongest relationship with FYGPA
  - Superscoring minimized prediction error by number of retests

- Impact on Diversity
  - Superscoring did not result in a less diverse admitted class
Impact of Superscoring on Subgroup Differences

Krista Mattern and Justine Radunzel
Research Question

Given that underserved students are less likely to retest as compared to their affluent peers, we wanted to investigate the extent to which superscoring increases, decreases, or has no impact on subgroup differences.
Current Study

Sample
2018 ACT-tested graduating class

Method
Compared subgroup differences in average ACT composite score for two scoring methods (most recent vs. superscoring)
  • Gender, race/ethnicity, household income, parental education

Results
We report subgroup differences using 2 metrics:
  • Mean differences or unstandardized differences (USTD): the difference between the mean value in two groups
  • Standardized differences (STD): the difference between the mean value in two groups, divided by the overall standard deviation.
Subgroup Unstandardized (USTD) and Standardized (STD) Differences in ACT Composite Scores by Scoring Method

<table>
<thead>
<tr>
<th>Group</th>
<th>Most recent score</th>
<th>Superscore</th>
<th>Most recent score - Superscore</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>USTD</td>
<td>STD</td>
</tr>
<tr>
<td>Annual family income group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than $36,000 (Low)</td>
<td>18.2</td>
<td>-5.7</td>
<td>-0.98</td>
</tr>
<tr>
<td>$36,000 to $80,000 (Mid)</td>
<td>20.7</td>
<td>-3.2</td>
<td>-0.55</td>
</tr>
<tr>
<td>Missing</td>
<td>20.0</td>
<td>-3.9</td>
<td>-0.67</td>
</tr>
<tr>
<td>More than $80,000 (High)</td>
<td>23.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent's Ed Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No college</td>
<td>17.9</td>
<td>-7.0</td>
<td>-1.21</td>
</tr>
<tr>
<td>Some college</td>
<td>19.8</td>
<td>-5.1</td>
<td>-0.88</td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>22.5</td>
<td>-2.4</td>
<td>-0.41</td>
</tr>
<tr>
<td>Missing</td>
<td>18.9</td>
<td>-6.0</td>
<td>-1.03</td>
</tr>
<tr>
<td>Beyond bachelors</td>
<td>24.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

On average, USTDs are 0.17 larger (on a 1 to 36 scale) for superscores as compared to the most recent scores

- In terms of STDs, superscoring increases subgroup differences by 0.02
Mean ACT Composite Scores by Scoring Method and Number of Times Tested

<table>
<thead>
<tr>
<th>Number of times tested</th>
<th>N (% )</th>
<th>Most recent score</th>
<th>Superscore</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>1,064,222 (56%)</td>
<td>19.3</td>
<td>5.5</td>
</tr>
<tr>
<td>2</td>
<td>465,650 (24%)</td>
<td>22.0</td>
<td>5.7</td>
</tr>
<tr>
<td>3</td>
<td>215,527 (11%)</td>
<td>23.3</td>
<td>5.5</td>
</tr>
<tr>
<td>4 or more</td>
<td>169,415 (9%)</td>
<td>23.9</td>
<td>5.1</td>
</tr>
<tr>
<td>Total</td>
<td>1,914,814(100%)</td>
<td>20.8</td>
<td>5.8</td>
</tr>
</tbody>
</table>

- 44% of the sample retested
- Students who tested more often tended to have higher ACT composite scores
- Differences in average ACT composite score (recent vs. superscore) ↑ as the number of testing occasions ↑
Scoring Method Differences in USTD and STD in ACT Composite Score by Number of Times Tested

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of Times Tested</th>
<th>USTD</th>
<th>STD</th>
<th>USTD</th>
<th>STD</th>
<th>USTD</th>
<th>STD</th>
<th>USTD</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One</td>
<td>Two</td>
<td>Three</td>
<td>Four or more</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual family income group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than $36,000 (Low)</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.02</td>
<td>-0.20</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>$36,000 to $80,000 (Mid)</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>-0.01</td>
<td>-0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.20</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>More than $80,000 (High)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent's Ed Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No college</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td>-0.20</td>
<td>0.01</td>
<td>-0.10</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.10</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.10</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Beyond bachelors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Unstandardized and standardized subgroup differences of most recent score - superscore.

Results by number of testing occasions:
- For students who tested once, most recent = superscore
- For students who tested twice, USTDs are 0.09 smaller and STDs are the same (0.00)
- For students who tested three times, USTDs are 0.06 smaller and STDs are 0.01 higher
- For students who tested 4+ times, USTDs are 0.09 smaller and STDs are 0.01 higher
Implications

• Subgroup differences are largely unaffected by the two scoring policies examined

• Slight increases in USTDs and STDs can be attributed to differences in retest rates among subgroups.
  • Controlling for the number of retests indicated that subgroup differences were more likely to decrease rather than increase with superscoring

• Broader awareness of ACT fee waiver policy for low-income students may help promote retesting among underserved students and reduce subgroup differences

• Caveat: Despite these positive findings, the results may change if retesting behavior changes significantly in the future in terms of who retests and how often.
Questions
Superscoring: College Policy and Usage