Assessment and Placement: A Summary of Recent CCRC Studies

Nikki Edgecombe
Community College Research Center
Teachers College, Columbia University

City College of San Francisco
April 3, 2012
2012 Assessment Studies

• *Do High-Stakes Placement Exams Predict College Success?* (CCRC Working Paper No. 41)
  – by Judith Scott-Clayton

• *Predicting Success in College: The Importance of Placement Tests and High School Transcripts* (CCRC Working Paper No. 42)
  – by Clive R. Belfield and Peter M. Crosta

Both papers are available on the CCRC website
The Highest-Stakes Test You’ve Never Heard Of?

- Vast majority (92 percent) of community colleges use placement exams to screen entrants for access to college-level coursework
- Majority of tested students are referred to remediation as a result
- Yet very little research has been done regarding the validity of these exams as a screening tool
- Students themselves often unaware of the importance of these exams (Venezia, Bracco, & Nodine, 2010)
Guidance for Interpretation of Results

# 1 – Studies use rigorous methodology to generate *predictive models* not experiments

# 2 – Predictive validity studies provide information about the degree to which the *use* of placement tests *may be* effective or ineffective (not on the test itself)

# 3 – Overall validity depends on the *effectiveness of the treatment* to which students are assigned based on placement test results (and the available evidence regarding the impact of remediation is mixed)
Goals of Scott-Clayton Analysis

• To describe the relationship between placement exam scores and grades in relevant college-level courses
• To better understand the practical tradeoffs involved in raising or lowering cutoffs
• To explore whether the use of additional measures could improve placement outcomes
Overview of Analysis

• Data:
  – 4 cohorts of degree-seeking entrants at Large, Urban Community College System (LUCCS)
  – 35,000 to 70,000 observations depending on analysis

• Success criteria:
  – Grades in first college-level course in relevant subject
  – Pass/fail, C or higher, B or higher criteria

• Measures of predictive validity:
  – Correlation coefficients/R-squareds (not included in this presentation but discussed at length in paper)
  – Placement accuracy/error rates
What is “placement accuracy”?

<table>
<thead>
<tr>
<th></th>
<th>Would succeed at college-level</th>
<th>Would not succeed at college-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placed into remediation</td>
<td><strong>Under-placed</strong></td>
<td><strong>Accurately placed</strong></td>
</tr>
<tr>
<td>Placed into college-level</td>
<td><strong>Accurately placed</strong></td>
<td><strong>Over-placed</strong></td>
</tr>
</tbody>
</table>

- We can’t directly observe potential outcomes in the top row
- But we can:
  1. Estimate relationship between test scores and outcomes for those placed directly into college level using logistic regression, then
  2. Predict outcomes for those placed into remediation
  3. Use predicted outcomes to simulate overall accuracy and error rates under different placement rules
- Extrapolation is a concern, BUT…
  - Limiting analysis to those near assignment cutoff does not change the story
  - Analysis is certainly relevant given policy trend towards *increasing* cutoffs
Figure 5. Probability of Gatekeeper Success, By Math Part 2 Scores

- T1 = Type I error: Not in DE but should be (or overplacement)
- T2 = Type II error: In DE but should not be (or underplacement)
Given the test in use, are cutoffs in the right place?

Predicted Severe Error Rates and Other Measures for Students Just Above and Just Below Placement Test Cutoffs

<table>
<thead>
<tr>
<th>Math (restricted to students +/- 5 points around algebra cutoff)</th>
<th>Placement Test Scores Only</th>
<th>All Students in Dev Ed</th>
<th>All Students in College Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe error rate</td>
<td>0.295</td>
<td>0.318</td>
<td>0.342</td>
</tr>
<tr>
<td>Severe overplacement rate</td>
<td>0.093</td>
<td>0.000</td>
<td>0.342</td>
</tr>
<tr>
<td>Severe underplacement rate</td>
<td>0.202</td>
<td>0.318</td>
<td>0.000</td>
</tr>
<tr>
<td>Remediation rate</td>
<td>0.703</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>College-level success rate (C or above) for those assigned to college level</td>
<td>0.556</td>
<td>n/a</td>
<td>0.517</td>
</tr>
<tr>
<td>Immediate college-level success rate for all those taking tests</td>
<td>0.165</td>
<td>0.000</td>
<td>0.517</td>
</tr>
</tbody>
</table>
Can we do better by incorporating other measures for placement?

Predicted Severe Error Rates and Other Measures Using Alternative Measures for Placement

<table>
<thead>
<tr>
<th>Measure (Math restricted to students +/- 5 points around algebra cutoff)</th>
<th>Placement Test Scores Only</th>
<th>Index of HS GPA/Units Only</th>
<th>Placement Test Scores + HS GPA/Units</th>
<th>Test Scores, HS GPA/Units + Local HS, Years Since HS</th>
<th>Use Students’ Best of Test Scores or HS Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe error rate</td>
<td>0.024</td>
<td>0.227</td>
<td>0.213</td>
<td>0.208</td>
<td>0.217</td>
</tr>
<tr>
<td>Severe overplacement rate</td>
<td>0.058</td>
<td>0.048</td>
<td>0.045</td>
<td>0.044</td>
<td>0.074</td>
</tr>
<tr>
<td>Severe underplacement rate</td>
<td>0.183</td>
<td>0.179</td>
<td>0.168</td>
<td>0.164</td>
<td>0.143</td>
</tr>
<tr>
<td>Remediation rate</td>
<td>0.748</td>
<td>0.747</td>
<td>0.747</td>
<td>0.747</td>
<td>0.666</td>
</tr>
<tr>
<td>College-level success rate (≥C) for those assigned to college level</td>
<td>0.670</td>
<td>0.708</td>
<td>0.734</td>
<td>0.747</td>
<td>0.676</td>
</tr>
<tr>
<td>Immediate college-level success rate for all those taking tests</td>
<td>0.169</td>
<td>0.179</td>
<td>0.185</td>
<td>0.189</td>
<td>0.226</td>
</tr>
</tbody>
</table>
Summary of Scott-Clayton Findings

• Placement test scores are more predictive in math than English
• System appears willing to tolerate high rates of under-placement in order to minimize over-placements
• Alternative measures do as good or better job at sorting students:
  – Using HS background measures alone to determine placement does as well in math and better in English than test scores alone
  – Combining test scores and HS measures could reduce severe placement errors by 12% while also improving college-level success rates by 10-15%.
  – But no combination of measures will eliminate mistakes completely
• Allowing students to test out of remediation based on “best of” either high school index or test score cutoff:
  – Would reduce remediation rates by 8 to 12 percentage points
  – While increasing college-level success rates
  – Generating large increases in percent succeeding at college-level in first semester
• Findings suggest that multiple measures should be used where possible and schools should retain flexibility to override score-based placements
Goals of Belfield & Crosta Analysis

• To assess the accuracy of these placement tests through replication of Scott-Clayton analysis
  – Measuring severe error rate through estimates of over-placement and under-placements

• To determine if incorporating school information into placement decisions improves accuracy

• To measure accuracy in predicting
  – College GPA / college credits earned
  – Probability pass college-level Math/English
Overview of Analysis

• Data:
  – Used several administrative data sets provided by a statewide community college system (SWCCS)
  – Multiple cohorts of students who took placement tests and enrolled in SWCCS between fall 2008 and summer 2010
  – Over 40 placement tests are used in state, but analysis draws from only 4
  – Sample sizes varied based on the analysis (and are noted where appropriate)

• Success criteria and measures of predictive validity are same as Scott-Clayton analysis
## Correlations Between Placement Test Scores and Grades in Developmental Classes

<table>
<thead>
<tr>
<th>Test</th>
<th>Math 1</th>
<th>Math 2</th>
<th>Math 3/4</th>
<th>English 1</th>
<th>English 2</th>
<th>Reading 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACCUPLACER</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algebra</td>
<td>0.12</td>
<td>0.07</td>
<td>0.03</td>
<td>0.03</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>0.13</td>
<td>0.13</td>
<td>0.08</td>
<td>0.08</td>
<td>0.13</td>
<td>0.12</td>
</tr>
<tr>
<td>Reading</td>
<td>0.20</td>
<td>0.15</td>
<td>0.14</td>
<td>0.10</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Sentence Skills</td>
<td>0.21</td>
<td>0.17</td>
<td>0.13</td>
<td>0.06</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>COMPASS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algebra</td>
<td>0.13</td>
<td>0.16</td>
<td>0.17</td>
<td>0.11</td>
<td>0.02</td>
<td>0.16</td>
</tr>
<tr>
<td>Pre-algebra</td>
<td>0.13</td>
<td>0.16</td>
<td>0.16</td>
<td>0.08</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Reading</td>
<td>0.25</td>
<td>0.17</td>
<td>0.12</td>
<td>0.17</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Writing</td>
<td>0.23</td>
<td>0.16</td>
<td>0.16</td>
<td>0.04</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>High School GPA</strong></td>
<td>0.36</td>
<td>0.36</td>
<td>0.35</td>
<td>0.34</td>
<td>0.35</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Interpretation of Correlations

• Overall correlations between placement test scores and developmental course grades are low
  – Example: For Math 1, the average correlation is 0.17 and for English 1, the average correlation is 0.06

• Placement test scores and grade associations are much stronger for developmental math courses than for developmental English/reading courses

• High school GPA is not only a better predictor but also more consistent than placement tests
**R-Squared Statistic for English Course Placements: Placement Test Scores and HS GPA**
(Sample restricted to students with HS data only)

<table>
<thead>
<tr>
<th></th>
<th>Placement Only</th>
<th>HS GPA Only</th>
<th>Placement + HS GPA</th>
<th>Placement + HS GPA + HS Variables</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACCUPLACER Sentence Skills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English ≥ B</td>
<td>0.004</td>
<td>0.105</td>
<td>0.105</td>
<td>0.112</td>
<td>2,270</td>
</tr>
<tr>
<td>English ≥ C</td>
<td>0.000</td>
<td>0.099</td>
<td>0.099</td>
<td>0.114</td>
<td>2,270</td>
</tr>
<tr>
<td>English ≥ D</td>
<td>0.000</td>
<td>0.087</td>
<td>0.087</td>
<td>0.103</td>
<td>2,270</td>
</tr>
<tr>
<td><strong>COMPASS Writing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English ≥ B</td>
<td>0.005</td>
<td>0.144</td>
<td>0.144</td>
<td>0.146</td>
<td>3,765</td>
</tr>
<tr>
<td>English ≥ C</td>
<td>0.003</td>
<td>0.116</td>
<td>0.116</td>
<td>0.119</td>
<td>3,765</td>
</tr>
<tr>
<td>English ≥ D</td>
<td>0.002</td>
<td>0.088</td>
<td>0.088</td>
<td>0.092</td>
<td>3,765</td>
</tr>
</tbody>
</table>

*R-squared values illustrate proportion of outcome variable (e.g., earning a B or higher in English) explained by various inputs (e.g., HS GPA)*
Interpretation of $R$-Squared Statistics

- Placement scores explain almost none—and in some cases actually none—of the variation in college-level grades.
- Placement test $R$-squared values are higher for math than for English.
- Placement tests are better predictors when the success criterion is more stringent ($\geq B$).
- High school GPA is by far a better predictor of success than the placement test.
Summary of Belfield & Crosta Findings

- Placement test scores are not especially good predictors of course grades in dev ed or other measures of performance (including success in gatekeeper classes).
- High school GPA is an extremely good and consistent predictor of college performance.
  - Challenge is this information may not be available for many students.
- Placement tests are associated with severe error rates.
  - 3 of every 10 test takers is either assigned to dev ed, despite being predicted to get at least a B in college-level English or assigned to college-level despite being predicted to fail the course.
  - HS GPA yields error rates half as big.
Questions for Discussion

• What are the implications for students and colleges of under-placement versus over-placement?

• Under what circumstances could it be justifiable to waive placement testing— and in so waive developmental education—for students who have HS GPAs above specified thresholds?
For more information:

Please visit us on the web at http://ccrc.tc.columbia.edu, where you can download presentations, reports, CCRC Briefs, and sign-up for news announcements.

Community College Research Center
Institute on Education and the Economy, Teachers College, Columbia University
525 West 120th Street, Box 174, New York, NY 10027
E-mail: ccrc@columbia.edu
Telephone: 212.678.3091

CCRC is funded in part by: Alfred P. Sloan foundation, Bill & Melinda Gates Foundation, Lumina Foundation for Education, The Ford Foundation, National Science Foundation (NSF), Institute of Education Sciences of the U.S. Department of Education