Using Construct Maps in Evidence Based Standard Setting

Denny Way, Julie Miles, Katie McClarty
Pearson
June 20, 2013

Presentation at the National Conference on Student Assessment, National Harbor, MD
Construct Maps

• Construct Maps allow standard-setting participants to use item response maps as a model of what a student at a given level knows and can do.

• The map is represented in software that allows standard-setting committee members to find out about the details of student performance at any given proficiency level, and to assist them in deciding where the cutoffs between performance levels should be.

Wilson & Carstensen (2007)
Example of a cut score setting map from:


<table>
<thead>
<tr>
<th>Scale</th>
<th>Multiple Choice</th>
<th>WR 1</th>
<th>WR 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>620</td>
<td>37</td>
<td>.30</td>
<td>2.3</td>
</tr>
<tr>
<td>610</td>
<td>15</td>
<td>.34</td>
<td>.26</td>
</tr>
<tr>
<td>600</td>
<td>28</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>590</td>
<td>27</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>580</td>
<td>19</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>570</td>
<td>34</td>
<td>.50</td>
<td>1.3</td>
</tr>
<tr>
<td>560</td>
<td>43</td>
<td>.53</td>
<td>.40</td>
</tr>
<tr>
<td>550</td>
<td>45</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>540</td>
<td>17</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>530</td>
<td>32</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td>520</td>
<td>33</td>
<td>.64</td>
<td></td>
</tr>
<tr>
<td>510</td>
<td>44</td>
<td>.67</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>18</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>490</td>
<td>20</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>480</td>
<td>40</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>470</td>
<td>47</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>460</td>
<td>10</td>
<td>2.2</td>
<td>.56</td>
</tr>
<tr>
<td>450</td>
<td>13</td>
<td>1.2</td>
<td>.40</td>
</tr>
<tr>
<td>440</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>430</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>410</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>390</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>380</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>360</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>350</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Construct Maps: A Good Idea

• Construct maps and related concepts have a rich history in standard setting and score interpretation.

• Typically, these maps are developed at the level of items on a test.

• This presentation provides an example of using a construct map at a higher level, specifically in organizing validity evidence collected in support of standard setting.
Setting Performance Standards

• Judgmental process of experts

• Historically content driven

• Basic, Proficient, Advanced

• In the NCLB era, content driven approach resulted in different standards for different states

• The common core emphasis on college and career readiness requires a different approach
Evidence Based Standard Setting

• Integrates content-centered judgments

• With the best available evidence from systematic research
Evidence Based Standard Setting Process

1. Define the outcome of interest
2. Develop research, data collection, and analysis plans
3. Synthesize the research results
4. Conduct standard-setting meeting with panelists
5. Continue to gather evidence
ADP Algebra II: College Readiness Definition

- Prepared for first-year credit bearing college mathematics courses
- Student should ultimately earn a B or better in the course without prior remediation
- College Algebra and Pre-Calculus
- 2-year, “typical” 4-year, and “more selective” 4-year institutions
ADP Algebra II: Research Studies

• Concurrent Studies
  – State exams (HI, IN, KY, NJ, PA, RI)
  – National exams (SAT, ACT, PSAT)

• Predictive Studies
  – College students across institutions and courses

• Content Judgment Studies
  – 133 faculty across institutions and courses
ADP Algebra II: Data Analysis and Presentation

Abbreviations
AL = Algebra
PC = Pre-Calculus
CC = Community College
4T = 4-year Typical Admittance Rate Institution
4S = 4-year More Selective Admittance Rate Institution

ACT-Exp. & Pred. Score
ACT-Concordance
Growth in 4T PC
B or better in 4T AL
B or better in 4T PC
B or better in 4S AL & CC PC
B or better in CC AL

Studies Key
National Concurrent
Predictive Study
Judgment Studies
ADP Algebra II: Panelist Meeting

- Briefing Book
  - Assessment background
  - Studies: content and data analyses
  - Crosswalk of study results

- See www.pearsonassessments.com/adpbriefingbook

- Panelists completed 3 rounds of judgments

- Round 3 recommendations were presented to policymaking body
Benefits of EBSS

- Performance standards informed by the best available data linking performance on high school assessments to performance in college

- Performance standards vertically aligned as a comprehensive system

- Data to support the reasonableness of the resulting performance standards

- More information available to provide stakeholders

- Flexible ways in which empirical data can be used in conjunction with content-based judgments
Challenges of EBSS

- Sufficient Evidence
- Presentation of Data
- Greater Complexity than Policymakers may Want
References


Keng, Murphy, & Gaertner (2012, April). *Supported by data: A comprehensive approach for building empirical evidence for standard setting*. Paper presented at NCME.


Contact

Denny Way, Ph.D.
Senior Vice President
Assessment & Instruction
Denny.way@pearson.com

Julie Miles, Ph.D.
Vice President
Assessment & Instruction
Julie.miles@pearson.com

Katie McClarty, Ph.D.
Director,
Center for College & Career Success
katie.mcclarty@pearson.com

http://researchnetwork.pearson.com