Using the ACGME Core Competencies to Evaluate Faculty
Sara Multerer, MD; Karen Hughes Miller, PhD; Michael Rowland, PhD; Kimberly Boland, MD; Charles Woods, MD
Departments of Pediatrics, University of Louisville School of Medicine, Louisville, Kentucky

Introduction
The Accreditation Council on Graduate Medical Education (ACGME) core competencies have been the cornerstone of evaluation of young physicians in training for the 21st century. However, to date there are no published faculty evaluation tools that use these same competencies. We present the first published evaluation of faculty using the core competencies to assess clinical faculty as teachers.

Methods & Materials
Quantitative Methods
- A 17-item faculty evaluation was created using the six ACGME core competencies as a framework.
- Responses to each evaluation item used a five point Likert scale (one = unsatisfactory, five = excellent).
- Data were collected using the New Innovations system and captured for calendar years 2010-2012.
- Cronbach’s α was used to assess the validity of the 17 items as a scale for all evaluations and for averaged item scores for individual faculty members.
- Results of incomplete evaluations (missing scores on ≥1 of 17 items) were compared to complete evaluations, adjusting for the number of items completed.
- Trends over time were assessed with Mantel-Haenszel chi square.

Qualitative Methods
- Comments were de-identified and sorted by faculty, department, and year.
- Using Pandit’s 1 variation of Glaser and Strauss’ Constant Comparison, comments were read and re-read by independent reviewers who coded the comments.
- Comments were then divided into two groups:
  • One for faculty receiving a “Top 20” teaching award at least two out of four years
  • And one for faculty receiving only one or no “Top 20” teaching awards.
- The analysis was then repeated.

Results
Evaluation of Completion Rates and Mean Scores per Scored Item During the Study Period

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluations, N</td>
<td>7079</td>
<td>2261</td>
<td>2496</td>
<td>2322</td>
<td></td>
</tr>
<tr>
<td>17 Items*, n (%)</td>
<td>5616 (79.3)</td>
<td>1904 (64.2)</td>
<td>1977 (79.2)</td>
<td>1785 (74.4)</td>
<td>&lt;.001* &lt;.001*</td>
</tr>
<tr>
<td>15-16 Items*, n (%)</td>
<td>1147 (16.2)</td>
<td>323 (43.4)</td>
<td>331 (13.3)</td>
<td>493 (21.2)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>1-14 Items*, n (%)</td>
<td>234 (3.3)</td>
<td>32 (1.4)</td>
<td>108 (4.3)</td>
<td>94 (4.0)</td>
<td></td>
</tr>
<tr>
<td>0 Items*, n (%)</td>
<td>82 (1.2)</td>
<td>2 (0.1)</td>
<td>80 (3.2)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Score Per Item, ANY Item(s) completed, Mean ± S.D.</td>
<td>4.63 ± 5.5</td>
<td>4.59 ± 5.5</td>
<td>4.64 ± 5.3</td>
<td>4.66 ± 5.2</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Score Per Item, ALL Items completed, Mean ± S.D.</td>
<td>4.67 ± 5.0</td>
<td>4.63 ± 5.3</td>
<td>4.67 ± 5.2</td>
<td>4.74 ± 5.4</td>
<td>&lt;.001*</td>
</tr>
</tbody>
</table>

17 Item Distribution into 6 Core Competency Domains

<table>
<thead>
<tr>
<th>Item Category</th>
<th>Patient Care</th>
<th>Medical Knowledge</th>
<th>Interpersonal &amp; Communication Skills</th>
<th>Professionalism</th>
<th>Practice-Based Learning and Improvement</th>
<th>Systems-Based Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items with “pure” domain</td>
<td>6</td>
<td>12*</td>
<td>7*</td>
<td>8*, 10, 12*</td>
<td>15*</td>
<td>18*, 175</td>
</tr>
<tr>
<td>Items with potential domain</td>
<td>1, 2, 5, 16*</td>
<td>2, 3, 4, 14*</td>
<td>1, 3, 4</td>
<td>14</td>
<td>13*</td>
<td></td>
</tr>
<tr>
<td>* Listed in a grouping labeled Patient Care and Professions on the evaluation form.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>† Listed in a grouping labeled Medical Knowledge on the evaluation form.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>§ Listed in a grouping labeled Practice-Based Learning and Improvement on the evaluation form.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‡ Listed in a grouping labeled Systems-Based Practice on the evaluation form.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analysis
- There were 7079 total resident evaluations (5816 complete, 79.3%) of 233 faculty members over the 3 yr. period. The mean score per item was 4.68 ± 5.0. Cronbach α was 0.98. The mean scale score per faculty member was 79.5 with variance of 26.5. Mean score per item was 4.67. Cronbach α was 0.99. Completion rate of all items declined from 84.2% to 74.7% (P<.001). Mean scores of incomplete evaluations adjusted for the number of items completed were lower than mean scores of complete evaluations across all time periods.
- Over 7,000 comments were collected from pediatrics residents’ evaluations of approximately 250 faculty members from 2008 to 2012. There were 22 “Top 20” faculty with 1318 total comments from this group.

Discussion
- This 17-item evaluation appears to have excellent validity for assessment of faculty performance in teaching, as our pediatrics residents perceive faculty quality based on these ACGME core-competency based items.
- Our residents rate their faculty overall very highly, which is consistent with the few prior studies in this area.
- Lower mean scores for incomplete evaluations, while still high overall, suggest that learners may not complete evaluation items for faculty they perceive as less effective teachers or role models.
- Qualitative analysis of resident evaluation comments reveals recognizable differences between award winning and regular faculty, helping to validate an instrument that tends to generate consistently high scores.
- The list of “traits” provides insight into how residents articulate the traits of exemplary teaching and primarily includes personal behaviors and creating a supportive teaching environment rather than medical knowledge or technical skill.
- Limitations include:
  • Single center study
  • Overall very narrow range of numeric evaluation scores of faculty and very high numeric score ratings
  • Opportunities for improvement found within the evaluation

Qualitative Evaluation Consistent “Top 20” Winners
Notable differences between “Top 20” and Total Group
- Far fewer negative scenarios
- Far fewer mentions of “lack of efficiency”.
- Frequent use of comparative language, e.g. best, outstanding, excellent.
- Comments with rich descriptors.
- Frequent mentions of patient care, patient advocacy

Descriptions of personal behaviors included:
- Approachable
- Knowledgeable
- Easy to work with
- Polite
- Efficient
- Role model
- Respectful

Descriptions of teaching environment and techniques included:
- Non-threatening
- Gives us autonomy
- Inclusive
- Takes time
- Can sometimes be intimidating
- Incorporate teachable moments
- Charged us to read
- Teaches at all levels
- Simplified; broke it down
- Asks questions
- Patient care is top priority

References