Faculty and residents share a single view of the ideal clinical teacher. The ideal teacher can be described as having four roles: physician, supervisor, teacher, and person. This description encompasses teaching behaviors, attitudes and interpersonal skills. The objective of this study was to identify themes found in resident feedback when they comment on better or worse than average teaching. We hypothesized that specific behaviors and characteristics will be associated with better and worse-than-average teaching scores.

### Results

#### Background

- 9.8% below-average evaluations and 9.4% of above-average evaluations correctly predicted based on comments alone (Cronbach α = 0.95).
- High reliability (>0.7) for 5 of 15 positive themes and all negative themes.
- One positive theme (p5) excluded for very poor reliability.
- All positive themes but p10 had significant associations with above-average teaching scores using Fisher’s two-sided exact test.
- All negative themes but n3 had significant associations with below-average teaching scores using Fisher’s two-sided exact test.
- Positive themes p3, p4, p6, p9, p11, and p15 were independently associated with above-average evaluations using logistic regression analysis.
- Negative themes n1, n2, n4, n5, n6, n7, n10, and n12 were independently associated with below-average evaluations using logistic regression analysis.

#### Methods

- IRB approved (MGH) and IRB exempted (Michigan).
- All resident evaluations of anesthesia teaching faculty 1/1/2007-12/31/2008 included.
- For each resident, calculated mean teaching score given to faculty, extracted all evaluations ≥ ±20% or ≤ -15% mean.
- One investigator (KB) read comments and developed list of recurring themes (Table 1 and 2).
- Two investigators (BH, JC) were blinded to teaching score and determined presence/absence of recurring themes in each evaluation, and predicted whether teaching score was above- or below-average.
- They calculated inter-rater reliability (Cronbach α) and compared relationship of theme to dichotomized teaching score (above- or below-average) using Fisher’s two-sided exact test.
- Themes with low reliability were excluded.
- Dichotomous outcome (above- or below-average) regressed on independent themes using logistic regression model.

#### Results

- 9,786 evaluations submitted by 117 residents for 162 faculty reviewed.
- 527 evaluations were ≥ ±20% mean teaching score and had comments.
- 285 evaluations were ≤ -15% mean teaching score and had comments.
- 9 evaluations excluded for uninformative comments.
- 198 evaluations excluded for having no themes from Table 1 or 2.

### Discussion

The comments in our study convey diagnostic teaching information. The numeric teaching score has little value independent of the associated comments, at least for above- or below-average evaluations. Limitations include selection bias, and the possibility of reporting bias. This study is predicated on the idea that the residents submitting evaluations accurately identify and report valid reasons for the high or low teaching score that they assign. This study provides an association between above-average teaching and the behaviors or traits in each of Ulian’s four domains that describe the ideal teacher: didactic, teacher, supervisor, and person. Themes are recast as recommendations in Table 3.

#### References


---

**Resident Characterization of Better and Worse than Average Clinical Teaching**

Bishr Haydar MD¹, Jonathan Charnin MD², Terri Voepel-Lewis MS RN¹, Keith Baker MD PhD²

¹Department of Anesthesiology, University of Michigan
²Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital

---

**Table 1: Positive Themes in High-Score Evaluations**

<table>
<thead>
<tr>
<th>Theme</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All resident evaluations of anesthesia teaching faculty</td>
<td>0.95</td>
</tr>
<tr>
<td>2. High reliability (&gt;0.7) for 5 of 15 positive themes and all negative themes</td>
<td></td>
</tr>
<tr>
<td>3. One positive theme (p5) excluded for very poor reliability</td>
<td></td>
</tr>
<tr>
<td>4. All positive themes but p10 had significant associations with above-average teaching scores using Fisher’s two-sided exact test</td>
<td></td>
</tr>
<tr>
<td>5. All negative themes but n3 had significant associations with below-average teaching scores using Fisher’s two-sided exact test</td>
<td></td>
</tr>
<tr>
<td>6. Positive themes p3, p4, p6, p9, p11, and p15 were independently associated with above-average evaluations using logistic regression analysis</td>
<td></td>
</tr>
<tr>
<td>7. Negative themes n1, n2, n4, n5, n6, n7, n10, and n12 were independently associated with below-average evaluations using logistic regression analysis</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Negative Themes in Low-Score Evaluations**

<table>
<thead>
<tr>
<th>Theme</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All resident evaluations of anesthesia teaching faculty</td>
<td>0.95</td>
</tr>
<tr>
<td>2. High reliability (&gt;0.7) for 5 of 15 positive themes and all negative themes</td>
<td></td>
</tr>
<tr>
<td>3. One positive theme (p5) excluded for very poor reliability</td>
<td></td>
</tr>
<tr>
<td>4. All positive themes but p10 had significant associations with above-average teaching scores using Fisher’s two-sided exact test</td>
<td></td>
</tr>
<tr>
<td>5. All negative themes but n3 had significant associations with below-average teaching scores using Fisher’s two-sided exact test</td>
<td></td>
</tr>
<tr>
<td>6. Positive themes p3, p4, p6, p9, p11, and p15 were independently associated with above-average evaluations using logistic regression analysis</td>
<td></td>
</tr>
<tr>
<td>7. Negative themes n1, n2, n4, n5, n6, n7, n10, and n12 were independently associated with below-average evaluations using logistic regression analysis</td>
<td></td>
</tr>
</tbody>
</table>

---

C.S. Mott
Children’s Hospital