Behaviors of Highly Professional Resident Physicians

Darcy A. Reed; Colin P. West; Paul S. Mueller; et al.


http://jama.ama-assn.org/cgi/content/full/300/11/1326

Correction
Contact me if this article is corrected.

Citations
This article has been cited 1 time.
Contact me when this article is cited.

Topic collections
Medical Practice; Medical Education; Quality of Care; Quality of Care, Other
Contact me when new articles are published in these topic areas.

Related Letters
Measuring Professionalism in Resident Physicians
Michael J. Peeters et al. JAMA. 2009;301(7):723.

In Reply:
Darcy A. Reed et al. JAMA. 2009;301(7):723.
Behaviors of Highly Professional Resident Physicians

Darcy A. Reed, MD, MPH
Colin P. West, MD, PhD
Paul S. Mueller, MD, MPH
Robert D. Ficalora, MD
Gregory J. Engstler
Thomas J. Beckman, MD

Educators have been criticized for not teaching and rigorously assessing the core values of medicine that determine “professionalism.” Nevertheless, demonstrating professionalism is expected of all physicians across the education continuum. A lack of conscientiousness among medical students correlates with negative faculty assessments of professionalism, and unprofessional behaviors in medical school predict disciplinary action by state medical boards. Although research has identified correlates of unprofessional behavior among students, few studies have examined professionalism during residency.

It is necessary to be able to measure professionalism before expecting learners to acquire it. Stern proposed that effective professionalism assessments should use multiple methods by numerous assessors over time, be based on observations within realistic contexts, involve conflict (situations likely to challenge professionalism), not be overly stringent (because professionals cannot be expected always to behave perfectly), be transparent so that learners understand whether assessments are formative or summative, and be symmetric such that everyone in the organization is evaluated similarly. The hallmark of these criteria is the triangulation of assessments by multiple observers in realistic contexts over time to obtain a complete and accurate picture of professionalism.

In this study, we used criteria suggested by Stern to assess professionalism among residents by combining multiple 360-degree observation-based assessments from numerous observers (peers, senior residents, faculty, medical students, and nonphysician professionals) in clinical contexts over a year. We used a comparative design to identify behaviors that distinguished highly professional residents from their peers.

Context Unprofessional behaviors in medical school predict high stakes consequences for practicing physicians, yet little is known about specific behaviors associated with professionalism during residency.

Objective To identify behaviors that distinguish highly professional residents from their peers.


Main Outcome Measures Professionalism as determined by multiple observation-based assessments by peers, senior residents, faculty, medical students, and nonphysician professionals over 1 year. Highly professional residents were defined as those who received a total professionalism score at the 80th percentile or higher of observation-based assessments on a 5-point scale (1, needs improvement; 5, exceptional). They were compared with residents who received professionalism scores below the 80th percentile according to in-Training Examination (ITE) scores, Mini-Clinical Evaluation Exercise (mini-CEX) scores, conscientious behaviors (percentage of completed evaluations and conference attendance), and receipt of a warning or probation from the residency program.

Results The median total professionalism score among highly professional residents was 4.39 (interquartile range [IQR], 4.32-4.44) vs 4.07 (IQR, 3.91-4.17) among comparison residents. Highly professional residents achieved higher median scores on the ITE (65.5; IQR, 60.5-73.0 vs 63.0; IQR, 59.0-67.0; P = .03) and on the mini-CEX (3.95; IQR, 3.63-4.20 vs 3.69; IQR, 3.36-3.90; P = .002), and they completed a greater percentage of required evaluations (95.6%; IQR, 88.1%-99.0% vs 86.1%; IQR, 70.6%-95.0%; P < .001) compared with residents with lower professionalism scores. In multivariate analysis, a professionalism score in the top 20% of residents was independently associated with ITE scores (odds ratio [OR] per 1-point increase, 1.07; 95% confidence interval [CI], 1.01-1.14; P = .046), mini-CEX scores (OR, 4.64; 95% CI, 1.23-17.48; P = .02), and completion of evaluations (OR, 1.07; 95% CI, 1.01-1.13; P = .02). Six of the 8 residents who received a warning or probation had total professionalism scores in the bottom 20% of residents.

Conclusion Observation-based assessments of professionalism were associated with residents’ knowledge, clinical skills, and conscientious behaviors.

Author Affiliations: Divisions of Primary Care Internal Medicine (Dr Reed), General Internal Medicine (Drs West, Mueller, Ficalora, and Beckman), Biostatistics (Dr West), and Information Services (Mr Engstler), Mayo Clinic College of Medicine, Rochester, Minnesota.

Corresponding Author: Darcy A. Reed, MD, MPH, Division of Primary Care Internal Medicine, Department of Internal Medicine, Mayo Clinic College of Medicine, 200 First St SW, Rochester, MN 55905 (reed.darcy@mayo.edu).

On Call Section Editor: Robert M. Golub, MD, Senior Editor.

JAMA, 2008;300(11):1326-1333 www.jama.com
Our conceptual framework was that professionalism is a multidetermined, synthetic competency encompassing attitudes, knowledge, skills, and behaviors. Therefore, we hypothesized that highly professional residents would be more likely than their peers to exhibit conscientious behaviors and demonstrate higher academic and clinical achievement.

METHODS
Study Design and Participants
We conducted a comparative study of all 148 first-year categorical internal medicine residents at Mayo Clinic between July 1, 2004, and June 30, 2007, (3 consecutive classes of interns) to identify behaviors that distinguished highly professional residents from their peers. This study was deemed exempt from review by the Mayo Clinic institutional review board.

Assessment of Professionalism
We used characteristics suggested by Stern for effective assessments of professionalism to design a longitudinal assessment of professionalism among residents. The assessment of professionalism includes 360-degree observation-based scores from peers (fellow first-year internal medicine residents who worked with the resident for at least 30 days on an inpatient medical service), senior residents (third-year internal medicine residents who had supervised the first-year resident on an inpatient medical service for at least 30 days), faculty (attending physicians who supervised the resident on an inpatient medical service for at least 2 weeks), medical students (third- or fourth-year students who were supervised by the resident on an inpatient rotation for at least 2 weeks), and nonphysician professionals (registered nurses, licensed practical nurses, and clinical assistants who had worked with the resident in an inpatient or outpatient setting for a minimum of 2 weeks). Peers, senior residents, faculty, medical students, and nonphysician professionals rated residents’ professionalism on a numeric scale from 1 to 5 (1, needs improvement; 3, average resident; 5, exceptional) for the items shown in Table 1. We collected scores from peers, senior residents, faculty members, medical students, and nonphysician professionals for all 148 residents throughout their first year of internal medicine residency.

Validity and Internal Structure of Professionalism Assessment
Content validity of the professionalism assessment was demonstrated by comprehensively reviewing published reports on professionalism and by developing instrument items that reflect published domains of professionalism and best practices in professionalism measurement. Our method of assessing professionalism fulfills the criteria suggested by Stern by incorporating assessments from multiple raters in realistic settings that regularly involve conflict (inpatient wards) over time (1 year of residency). This measurement is not overly stringent because residents are observed and assessed on numerous occasions; therefore, residents may receive one or more poor evaluations yet maintain a positive overall assessment. Furthermore, the purpose of the assessment is transparent. Assessment data are used by residents for ongoing self-improvement. Residents have timely access to their confidential assessment data, and faculty advisors also review data with residents for the purpose of formative assessment. Assessment data from all observer groups except faculty are anonymous. Finally, the assessment is symmetric because all residents, faculty, students, and nonphysician professionals receive observation-based, 360-degree assessments.

The dimensionality of raw scores from items from each observer group was studied using exploratory factor analysis with oblique rotation. Factors with Eigenvalues of more than 1 were retained, and the final factor structure was confirmed by inspecting the corresponding scree plot. Items with factor loadings of more than 0.3 were retained. Internal consistency reliability of items comprising each factor was determined using Cronbach α, for which values of more than 0.7 were considered acceptable.

Definition of Comparison Groups
We calculated the mean of all individual assessments obtained (from peers, senior residents, faculty, students, and nonphysician professionals together) over a 1-year period to determine a numeric total professionalism score for each resident. Each individual observation was assigned equal weight in the analysis so that the total professionalism score was the simple mean of all individual assessments obtained. Scores were not averaged by observer group and were not adjusted for the number of evaluations obtained per resident.

To examine potential differences in results based on using weighted vs nonweighted total professionalism scores, we repeated the analyses using scores weighted by observer group. The weighted professionalism score was determined by calculating the mean score from each observer group and then calculating the mean of these means to form the total weighted professionalism score. If 1 group did not evaluate a resident at all, this group was removed from the numerator and the denominator. Using this weighted professionalism score did not alter the results; therefore, only unweighted results are presented.

Residents with total scores at the 80th percentile or higher of the sample (30 residents) comprised the highly professional group. Residents with total scores below the lower 80th percentile (118 residents) served as the comparison group. These comparison groups were determined a priori because the objective of this study was to identify behaviors that distinguished highly professional residents from their peers. We chose to focus on high professionalism rather than low or unprofessionalism, recognizing that the “unprofessional” label may be problematic because an individual’s professionalism varies by situation and context.

©2008 American Medical Association. All rights reserved.

(Reprinted) JAMA, September 17, 2008—Vol 300, No. 11 1327
Table 1. Professionalism Assessment for First-Year Internal Medicine Residents

<table>
<thead>
<tr>
<th>Observer Assessment</th>
<th>No. of Residents Observed</th>
<th>No. (%) of Observations</th>
<th>No. of Observations Per Resident Mean (SD)(a)</th>
<th>Score, Median (IQR)(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment</td>
<td>147</td>
<td>2027 (25.60)</td>
<td>13.8 (2.90)</td>
<td>4.19 (3.99-4.44)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>of cross-coverage</td>
<td></td>
<td></td>
<td>4.20 (4.00-4.44)</td>
<td></td>
</tr>
<tr>
<td>and on-call tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective communication with families, patients, and team</td>
<td>4.13 (3.90-4.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of integrity</td>
<td></td>
<td></td>
<td>4.30 (4.12-4.55)</td>
<td></td>
</tr>
<tr>
<td>Dedication to learning environment</td>
<td>4.10 (3.85-4.28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as future coworker or team member</td>
<td>4.35 (4.00-4.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as a physician for a family member</td>
<td>4.18 (3.86-4.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior resident</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment</td>
<td>145</td>
<td>1136 (14.40)</td>
<td>7.8 (1.70)</td>
<td>4.17 (3.92-4.41)</td>
</tr>
<tr>
<td>Level of integrity</td>
<td></td>
<td></td>
<td>4.28 (4.11-4.57)</td>
<td></td>
</tr>
<tr>
<td>Commitment to own education</td>
<td>4.16 (4.00-4.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition of own limitations</td>
<td>4.09 (3.75-4.28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as a future coworker or team member</td>
<td>4.28 (3.95-4.55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as a physician for a family member</td>
<td>4.11 (3.71-4.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment</td>
<td>148</td>
<td>4128 (52.10)</td>
<td>27.9 (5.70)</td>
<td>4.01 (3.80-4.15)</td>
</tr>
<tr>
<td>Humanistic qualities and professionalism</td>
<td>4.00 (3.72-4.26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commitment to own education</td>
<td>4.20 (4.04-4.31)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical student</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment</td>
<td>44</td>
<td>90 (1.10)</td>
<td>2.0 (0.70)</td>
<td>4.63 (4.08-4.81)</td>
</tr>
<tr>
<td>Level of integrity</td>
<td></td>
<td></td>
<td>4.83 (4.37-5.00)</td>
<td></td>
</tr>
<tr>
<td>Commitment to student’s education</td>
<td>4.50 (4.00-5.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as a future coworker or team member</td>
<td>4.58 (4.00-5.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desirability as a physician for a family member</td>
<td>4.50 (4.00-5.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonphysician professionals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall assessment</td>
<td>112</td>
<td>534 (6.80)</td>
<td>4.8 (2.10)</td>
<td>4.00 (3.75-4.17)</td>
</tr>
<tr>
<td>Responsive to concerns of nonphysician professionals</td>
<td>4.00 (3.75-4.26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstrates empathy and compassion for patients</td>
<td>4.00 (3.75-4.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treats nonphysician professionals with respect</td>
<td>4.13 (3.89-4.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completes patient care tasks in a timely manner</td>
<td>3.88 (3.67-4.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication with patients, families, and nonphysician professionals</td>
<td>3.88 (3.67-4.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: IQR, interquartile range.
\(a\)Means do not sum to total because the number of residents assessed among observer groups varied.
\(b\)Total professionalism score measured on a 5-point scale in which 1 indicates needs improvement; 3, average resident; and 5, exceptional resident.

Independent Variables

Independent variables were selected based on knowledge, skills, and behaviors believed to reflect various attributes of professionalism, as suggested by the American Board of Internal Medicine\(^10\) and others.\(^7,\)\(^11\) We examined “conscientious behaviors” using the percentage of evaluations completed by residents (number of required evaluations of peers, faculty, and rotations completed by residents divided by the number of evaluations assigned to residents) and residents’ conference attendance (number of required didactic sessions attended by residents). Conference attendance was measured using a card-swipe system wherein residents’ presence at conferences are recorded electronically using a bar code on a pocket card.

To measure residents’ medical knowledge, we examined residents’ scores on the In-Training Examination (ITE). The ITE is a standardized multiple-choice examination developed by the American College of Physicians and is administered annually to the majority of US internal medicine residency programs. Residents at Mayo Clinic take the examination annually in October. The ITE scores have been shown to be internally consistent, reliable measures of medical knowledge in internal medicine\(^18\) that correlate highly with American Board of Internal Medicine certification examination scores.\(^17\)

To measure residents’ clinical skills, we obtained Mini-Clinical Evaluation Exercise (mini-CEX) scores for residents. The mini-CEX is a focused history and physical examination of a real patient observed by a faculty physician. The mini-CEX has been shown to be an internally consistent, reliable, and feasible method for assessing residents’ clinical skills.\(^18,\)\(^19\) Faculty assessed residents on a 1 to 5 scale (1, needs improvement; 3, average resident; 5, exceptional) on medical interviewing skills, physical examination skills, clinical judgment, counseling skills, and organization and efficiency. Formal faculty assessment of
professionalism was not included in the mini-CEX evaluation. The mean of all mini-CEX scores per resident was used to determine an overall mini-CEX score for each resident.

We also recorded whether residents received an official warning of deficiency or probation from the residency program. Warnings and probation are issued by a committee of program faculty and leadership for both serious deficiencies in residents’ academic performance and breaches in medical or ethical conduct. We hypothesized that highly professional residents would be less likely than their peers to receive program warnings or probation.

Data Analysis

Due to small sample sizes within subgroups and expectations of skewed ordinal data, residents with total professionalism scores in the top quintile of the sample were compared with their peers using the Fisher exact test for binary variables and Wilcoxon rank sum test for continuous or ordinal variables. Wilcoxon signed rank tests for paired data were used to conduct comparisons of residents’ professionalism scores between groups of observers and between time frames (July-September vs April-June). We used the Spearman ρ to measure correlations between total professionalism scores rendered by the 5 groups of observers. We performed bivariate and multivariable logistic regression analyses to examine relationships between independent variables and total professionalism scores (top 20% vs lower 80%). Independent variables included residents’ conference attendance, percentage of required evaluations completed, ITE scores, and mini-CEX scores as continuous variables, and receipt of a warning of deficiency or probation from the residency program as a dichotomous variable (yes or no). Analyses were based on all available data; therefore, individuals were excluded only from analyses involving variables for which those individuals’ data were missing.

Stepwise forward selection was applied to model building. Variables were added to the multivariate model according to level of significance (P < .10) in bivariate analyses. Multivariate modeling using alternative variable inclusion thresholds between 0.05 and 0.20 did not alter the results. Model fit was examined using the Hosmer-Lemeshow goodness-of-fit test.30 Model variables were examined for evidence of collinearity using Spearman correlation coefficients, and we avoided simultaneously modeling variables with correlation coefficients of more than 0.6. For all analyses, P < .05 was considered statistically significant. Data were analyzed using STATA 8.0 (STATA Corp, College Station, Texas).

Additional analyses examining associations between independent variables and professionalism scores as a continuous outcome, and evaluating conference attendance in quartiles and deciles as categorical independent variables did not alter the results and are therefore not reported.

RESULTS

Resident Characteristics

The 148 first-year categorical internal medicine residents were graduates of 85 different medical schools from all 4 regions of the United States (42% central, 14% northeastern, 10% western, and 8% southern states) and 17 other countries (26%). Thirty-seven percent of residents were women compared with 43.1% of all residents in US internal medicine residency programs between 2004 and 2006 (P = .14), and 26% were international medical graduates compared with 43% nationally (P < .001).21-23 There were no differences between residents in this study and all first-year residents in US internal medicine programs according to proportion of Alpha Omega Alpha designees (17% vs 15%, P = .57).34

The median ITE score among 143 (96.6%) residents was 63.0 (interquartile range [IQR], 60.0-68.0), and the full range was 52 to 84 from a possible range of 0 to 100 points (Table 2). A median of 4.0 (IQR, 2.0-6.0) mini-CEX examinations per resident were conducted for 146 (98.6%) residents who achieved a median mini-CEX score of 3.72 (IQR, 3.40-4.0), range 1.8-5.0. Residents’ attendance at required di-

Table 2. Association of Variables With Professionalism Among First-Year Internal Medicine Residents

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>No. of Residents Observed</th>
<th>All Residents (n = 148)</th>
<th>Highly Professional Residents (n = 30)</th>
<th>Comparison Residents (n = 118)</th>
<th>P Value&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total professionalism score&lt;sup&gt;d&lt;/sup&gt;</td>
<td>148</td>
<td>4.14 (3.96-4.27)</td>
<td>4.39 (4.32-4.44)</td>
<td>4.07 (3.91-4.17)</td>
<td>N/A</td>
</tr>
<tr>
<td>In-training examination score&lt;sup&gt;e&lt;/sup&gt;</td>
<td>143</td>
<td>63.0 (60.0-68.0)</td>
<td>65.5 (60.5-73.0)</td>
<td>63.0 (59.0-67.0)</td>
<td>.03</td>
</tr>
<tr>
<td>Mini-CEX score&lt;sup&gt;f&lt;/sup&gt;</td>
<td>146</td>
<td>3.72 (3.40-4.0)</td>
<td>3.95 (3.63-4.20)</td>
<td>3.69 (3.36-3.90)</td>
<td>.002</td>
</tr>
<tr>
<td>No. of didactic conferences attended</td>
<td>145</td>
<td>66 (52-76)</td>
<td>69 (58-80)</td>
<td>65 (50-74)</td>
<td>.07</td>
</tr>
<tr>
<td>Evaluations completed, %</td>
<td>147</td>
<td>88.6 (77.0-95.8)</td>
<td>95.6 (88.1-99.0)</td>
<td>86.1 (70.6-95.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Program warning or probation, No. (%)</td>
<td>148</td>
<td>8 (5.4)</td>
<td>0</td>
<td>8 (5.4)</td>
<td>.36&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Abbreviations: Mini-CEX, Mini-Clinical Evaluation Exercise; N/A, not applicable.
<sup>a</sup>Residents with total professionalism scores in the top 20% of the sample.
<sup>b</sup>Residents with total professionalism scores in the lower 80% of the sample.
<sup>c</sup>Wilcoxon rank sum test except as noted.
<sup>d</sup>Scored on 5-point scale for which 1 indicates needs improvement; 3, average resident; 5, exceptional resident.
<sup>e</sup>Range of possible scores 0 to 100.
<sup>f</sup>Fisher exact test.
dactic conferences ranged from 6 to 118; the median number of conferences attended was 66 (IQR, 52-76). Residents completed the majority (median, 88.6% [IQR, 77.0%-95.8%]) of required evaluations (of peers, faculty, rotations) assigned to them. Conference attendance and evaluation completion rate data were obtained for 145 (98.0%) and 147 (99.3%) of residents, respectively. Six residents received an official warning of deficiency from the residency program and another 2 residents were placed on probation. Residents in this sample accessed their individual assessment data electronically a median of 19 (IQR, 1-42) times. Thirty residents (20.3%) never accessed their assessment data during the 1-year study period.

Professionalism Among Residents

Professionalism assessment data were obtained for all 148 (100%) residents. A total of 7915 observation-based assessments were obtained from observer groups (Table 1). The mean (SD) total assessments obtained per resident was 53.5 (10.2). A total of 91.9% of peers, 96.7% of senior residents, 82.6% of faculty, 92.9% of medical students, and 87.8% of nonphysician professionals participated in the assessments.

Total professionalism scores among the 148 first-year residents ranged from 2.39 to 4.57 with a median of 4.14 (IQR, 3.96-4.27; Table 1). Total professionalism assessments of residents were similar among peers (4.19, IQR, 3.99-4.44) and senior residents (4.17; IQR, 3.92-4.41), but peers and senior residents gave significantly higher scores than faculty (4.01; IQR, 3.80-4.15) and nonphysician professionals (4.00, IQR, 3.75-4.17; all P < .001). Medical students’ median overall rating of residents’ professionalism (4.63, IQR, 4.08-4.81) was higher than were peers, senior residents, faculty, and nonphysician professionals (all P < .001). There was modest correlation between total professionalism scores rendered by peers and senior residents (Spearman r = 0.47; P = .001), and by senior residents and faculty (Spearman r = 0.44; P = .003); however, there were no significant correlations between professionalism scores rendered by medical students or nonphysician professionals and any of the other groups of observers. There were no significant differences in median professionalism scores in the first 3 months of residency (July-September, 4.14; IQR, 3.97-4.51) compared with the last 3 months of the first year of residency (April-June, 4.20; IQR, 4.0-4.45; P = .65).

### Internal Structure of Professionalism Assessment

Factor analysis revealed item clusters comprising 5 factors (Eigenvalue); peer (7.2), senior resident (4.2), medical student (3.4), nonphysician professionals (2.5), and faculty (1.3). The percent variance explained by the variables comprising this model was 92%. Cronbach α for the combined item scores was 0.91. Cronbach α for peer was 0.95; senior resident, 0.88; medical student, 0.88; nonphysician professionals, 0.95; and faculty, 0.28.

### Factors Associated With Professionalism Among Residents

The median total professionalism score among highly professional residents was 4.39 (IQR, 4.32-4.44) compared with 4.07 (IQR, 3.91-4.17) among the comparison group. Residents with professionalism scores in the top 20% achieved higher median scores on the ITE (65.5; IQR, 60.5-73.0 vs 63.0; IQR, 59.0-67.0; P = .03) and mini-CEX (3.95; IQR, 3.63-4.20 vs 3.69; IQR, 3.36-3.90; P = .002) compared with residents with lower professionalism scores (Table 2). Residents with high professionalism scores also completed a greater percentage of required evaluations of peers, faculty, and rotations than residents with lower scores (95.6%; IQR, 88.1%-99.0% vs 86.1%; IQR, 70.6%-95.0%; P < .001). None of the 8 residents who received a warning of deficiency or probation received professionalism scores in the top 20%. Although there was no statistically significant association between receipt of a warning or probation and total professionalism scores in the top 20%, further inspection of the data revealed that 6 of the 8 residents who received a warning or probation had total professionalism scores in the bottom 20% of residents.

The median total professionalism scores among the 30 residents who never accessed their individual assessment data during the study period was 4.17 (IQR, 3.94-4.41) compared with 4.13 (IQR, 3.96-4.25) among the 118 residents who accessed their assessment data at least once (P = .11). There were no significant differences in professionalism scores received during the first 3 months compared with the last 3 months of internship among residents who never accessed their assessment data and those who viewed their data at least once.

Table 3 shows the results of bivariate and multivariate logistic regres-

---

**Table 3. Logistic Regression Analysis of Factors Associated With High Professionalism Among Internal Medicine Residents**

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Bivariate Analysis</th>
<th>Multivariate Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>In-training examination score</td>
<td>1.08 (1.02-1.15)</td>
<td>.01</td>
</tr>
<tr>
<td>Mini-CEX score</td>
<td>5.49 (1.71-17.57)</td>
<td>.004</td>
</tr>
<tr>
<td>Didactic conferences attended</td>
<td>1.03 (1.00-1.06)</td>
<td>.04</td>
</tr>
<tr>
<td>Evaluations completed, %</td>
<td>1.08 (1.03-1.14)</td>
<td>.003</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; Mini-CEX, Mini-Clinical Evaluation Exercise; OR, odds ratio.

1. Estimated OR of association between 1-point increase in independent variable and total professionalism score in the top 20% of sample.
2. Scored on a 0-1 point scale, 1 indicates needs improvement, 3, average resident; 5, exceptional resident.
3. Number of required didactic conferences attended.
4. Percentage of required evaluations of peers, faculty, and rotations completed.
sion for a 1-unit increase in independent variables. After multivariate adjustment, clinical skills as measured by mini-CEX scores were significantly associated with high professionalism scores (adjusted odds ratio [OR], 4.64; 95% confidence interval [CI], 1.23-17.48; P = .02). A 1-point increase in residents’ ITE scores was independently associated with total professionalism scores in the top 20% of residents (adjusted OR, 1.07; 95% CI, 1.01-1.14; P = .046). However, the range of possible scores on the ITE is 0 to 100 and the observed range in our sample was 52 to 84, so that a 10- or 15-point difference in ITE scores can occur and is educationally significant. A 10-point increase in ITE score was associated with an adjusted OR of 1.97 (95% CI, 1.11-3.62) for professionalism scores in the top quintile and a 15-point increase in ITE score was associated with an adjusted OR of 2.76 (95% CI, 1.16-6.89) for being in the highly professional group. Finally, the percentage of required evaluations completed by residents during their first year was associated with professionalism (adjusted OR, 1.07; 95% CI, 1.01-1.13; P = .02). Thus, a 20% increase in evaluation completion rate was associated with an adjusted OR of 3.87 (95% CI, 1.28-10.83) for professionalism scores in the top quintile. The educational significance of this result is that, for example, a resident with an evaluation completion rate of 70% has an odds nearly 4 times higher that he or she would have a total professionalism score among the top 20% than a resident with a 50% evaluation completion rate. There was no observed association between residents’ conference attendance and degree of professionalism in multivariate analysis.

COMMENT

This comparative study identifies specific behaviors, knowledge, and skills (completion of evaluations, ITE scores, and mini-CEX scores) that are associated with professionalism. Because these factors are measured at the majority of US internal medicine residency programs, these findings should apply broadly to professionalism assessment in graduate medical education.

The association between residents’ completion of required evaluations and professionalism scores is consistent with prior studies showing that certain “conscientious behaviors” in medical school predict professionalism later in training or in practice. Our study extends this finding to graduate medical education and suggests that evaluation completion rate may be an early surrogate for professionalism that can be measured in internship and used to monitor professionalism during residency training.

Many conceptual frameworks of professionalism include excellence in medical knowledge and clinical skills. The association between residents’ ITE scores and professionalism assessments demonstrated in this study, associations between medical school grades and professionalism scores in internship, and associations between Medical College Admission Test scores and disciplinary actions by state medical boards all support this framework. Additional studies are needed to further explore relationships between medical knowledge and professionalism.

Residents’ clinical skill, as measured by mini-CEX scores, was the variable most strongly associated with professionalism in this study. To our knowledge, this association has not been previously demonstrated and lends additional support for the concept of professionalism as a multidetermined competency. Although there is compelling evidence for the validity and reliability of the mini-CEX as a measure of clinical skill, mini-CEX assessments by faculty may be influenced by professional attributes of residents such as empathy and humanism demonstrated during clinical encounters. Although professionalism was not specifically scored in this mini-CEX, it may be expected that residents who are viewed as highly professional are also seen as clinically excellent. Future studies incorporating other measures of clinical skill such as unannounced standardized patients and simulation-based assessments may advance understanding of the relationship between clinical skills and professionalism.

Validity evidence supports our method for assessing professionalism among internal medicine residents. Content evidence for our assessment is based on an approach to professionalism suggested by Stern21 and Norcini22, which states that professional behaviors are best determined by seeking verification from numerous sources, which in our instrument includes assessments by peers, senior residents, faculty, students, and nonphysician professionals. Internal structure evidence is supported by a predicted, multidimensional assessment of residents’ professional behaviors and high internal consistency reliability. Criterion validity is demonstrated by significant associations between professionalism scores and residents’ knowledge, clinical skills, and professional behaviors.

Our professionalism assessment is also strengthened by including peer and nonphysician professional perspectives. Although peer assessments that fail to ensure anonymity are unlikely to be accurate, confidential and formative peer assessments (as was the case in this study) may provide valuable information about professionalism. Despite well-documented challenges in communication and collaboration between nonphysician professionals and physicians in clinical contexts, few professionalism studies have included nonphysician professionals’ views. Yet, nonphysician professionals are often positioned at the interface between residents and patients, affording them a unique opportunity to observe professional behaviors.

There are several limitations to this study. First, our assessment, although encompassing multiple domains, cannot assess every domain of professionalism and may underrepresent cer-
tain areas such as empathy and humanism.\textsuperscript{10,11} Second, although the professionalism assessment includes views from multiple groups, it did not include assessments by patients. Third, because this study was conducted at a single internal medicine residency program, the results may not be generalizable to residents at other programs or in other specialties. However, we have shown that the residents in this study are similar to all US internal medicine residents according to certain demographic variables. Furthermore, the independent variables measured in this study (evaluation completion rates, conference attendance, ITE scores, mini-CEX scores) are measured at the majority of US internal medicine residency programs so that the study findings should be meaningful to other programs.

Fourth, although our sample size was adequate to detect statistically and educationally significant differences between groups, a larger sample may have provided greater stability of the estimates. Fifth, items comprising our assessment existed as separate forms in the Mayo Clinic electronic environment, which may have increased the likelihood of discovering the identified factors. Yet item groupings overall and within most factors demonstrated excellent internal consistency reliability. Nevertheless, the Cronbach $\alpha$ was low for items comprising the faculty factor. This is likely because the magnitude of $\alpha$ is proportional to the number of items on a scale,\textsuperscript{14} and based on factor analysis, the faculty factor contained only 2 items.

Sixth, the professionalism scores are fairly high, indicating a possible ceiling effect in scores. However, we note that score inflation is well-documented among performance assessments in medicine\textsuperscript{34,35} and is not unique to our study. Despite this concern, there was adequate variation in professionalism scores to demonstrate meaningful differences between groups. Seventh, we used modified mini-CEX questions with a 5-point rather than 9-point scoring method and had a median of 4 examinations per resident, which may limit reliability, but likely provides sufficient precision for this assessment.\textsuperscript{36} Eighth, this study suggests associations of behavioral factors with professionalism but cannot confirm that these relationships are causal.

What are the implications of findings from this and similar studies? Educators believe that professionalism can be taught\textsuperscript{2} and most medical schools have formal professionalism curricula in place,\textsuperscript{37,38} but didactic instruction alone is insufficient to instill professionalism among trainees.\textsuperscript{1,39,40} Additional strategies such as explicit and consistent role modeling of professional behaviors, reflection, and self-assessment are needed to encourage the development of mindful, professional practitioners.\textsuperscript{41} In addition to teaching, personal and environmental factors (such as institutional culture, informal and formal curricula, and practice characteristics) influence professionalism,\textsuperscript{42} leading some authorities to suggest that institutional culture must evolve to encourage professionalism.\textsuperscript{39} This cultural transformation will be particularly challenging given heightened pressures on physicians to enhance productivity in an environment of increasing regulation and diminishing time to interact with learners.\textsuperscript{39,40,43} Nevertheless, a multifaceted approach to addressing learner behaviors and environmental factors is needed to promote professionalism.

The results of this comparative study further the understanding of professionalism in medicine. Our findings strengthen the notion of professionalism as a multidetermined construct by contributing new information about specific knowledge, skills, and behaviors associated with professionalism during residency. These factors can be measured within residency programs and may serve as a model for assessing professionalism among resident physicians.

Author Contributions: Dr Reed had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Reed, West, Mueller, Beckman. Acquisition of data: Reed, Ficler, Engster, Beckman. Analysis and interpretation of data: Reed, West, Mueller, Beckman. Drafting of the manuscript: Reed, Mueller, Beckman. Critical revision of the manuscript for important intellectual content: Reed, West, Mueller, Ficlera, Engster, Beckman. Statistical analysis: Reed, West, Beckman. Obtained funding: Reed. Administrative, technical, or material support: Mueller, Ficler, Engster, Beckman. Study supervision: Mueller, Beckman.

Financial Disclosures: None reported.

Funding/Support: Dr Reed received support from an Educational Innovations Award from the Mayo Clinic.

Role of the Sponsors: No funding organization or sponsor had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

REFERENCES