

# Assessing Professional Behavior: Yesterday, Today, and Tomorrow

Louise Arnold, PhD

## ABSTRACT

**Purpose.** The author interprets the state of the art of assessing professional behavior. She defines the concept of professionalism, reviews the psychometric properties of key approaches to assessing professionalism, conveys major findings that these approaches produced, and discusses recommendations to improve the assessment of professionalism.

**Method.** The author reviewed professionalism literature from the last 30 years that had been identified through database searches; included in conference proceedings, bibliographies, and reference lists; and suggested by experts. The cited literature largely came from peer-reviewed journals, represented themes or novel approaches, reported qualitative or quantitative data about measurement instruments, or described pragmatic or theoretical approaches to assessing professionalism.

**Results.** A circumscribed concept of professionalism is available to serve as a foundation for next steps in assessing professional behavior. The current array of assess-

ment tools is rich. However, their measurement properties should be strengthened. Accordingly, future research should explore rigorous qualitative techniques; refine quantitative assessments of competence, for example, through OSCEs; and evaluate separate elements of professionalism. It should test the hypothesis that assessment tools will be better if they define professionalism as behaviors expressive of value conflicts, investigate the resolution of these conflicts, and recognize the contextual nature of professional behaviors. Whether measurement tools should be tailored to the stage of a medical career and how the environment can support or sabotage the assessment of professional behavior are central issues.

**Final thought.** Without solid assessment tools, questions about the efficacy of approaches to educating learners about professional behavior will not be effectively answered.

*Acad. Med.* 2002;77:502–515.

## INTRODUCTION

Promoting professional behavior is currently a chief concern across the continuum of medical education. A critical component of this initiative involves assessment. It is through assessment that educators can gauge the progress

their learners are making in becoming and being professional; and it is through assessment that they can ascertain the success of programs promoting professionalism. Indeed, assessment of both learners and programs is a requirement that medical educators must now fulfill. The central role of assessment in promoting professionalism, then, requires an examination of the state of the art.

This article offers my interpretation of the literature on assessing professional behavior. I examined articles and books collected over the years through searches of Medline and ERIC from 1966 onward, CINAHL from 1982, and PsycINFO from 1984; cited in confer-

ence proceedings, bibliographies, and reference lists; and recommended by experts. Search terms included professionalism; professional behavior; specific elements of professionalism such as altruism, duty, and humanism; noncognitive characteristics and traits; physician role; patient–physician relationships; attitudes; personality; and medical education/medical students.

By considering assessment in its broadest sense,<sup>1</sup> I asked two key questions of the literature identified.

- What does this paper or book teach us about collecting information to make decisions regarding the profes-

*Dr. Arnold is professor of medicine and sociology and associate dean for medical education and research, University of Missouri–Kansas City School of Medicine, Kansas City, Missouri.*

*Correspondence and requests for reprints should be addressed to Dr. Arnold, University of Missouri–Kansas City School of Medicine, 2411 Holmes, Kansas City, MO 64108; telephone: (816) 235-1806; fax: (816) 235-5277.*

---

sional behaviors of learners or practitioners, educational processes, or programs?

- How does this study bear on current initiatives in medical education to find sound methods for evaluating professional behaviors in medical school, residency, and practice.

I used the following considerations as guides to choosing the literature to cite. I stressed studies from medical education but included work from related fields that offered fresh insights. I cited work that represented trends, as well as investigations that contained novel approaches. I emphasized reports about assessment instruments that included either quantitative or qualitative data about their measurement properties, but I also selected pragmatic descriptions of assessment as well as basic, theoretical, and conceptual investigations. By far, most of the literature cited was peer reviewed.

The literature prompted the following topics for exploration: the concept of professionalism, key measurement methods available for assessing professional behavior, and recommendations for improving the evaluation of professionalism. I address each of these topics in turn, and have organized the reference list according to the topics.

## THE CONCEPT OF PROFESSIONALISM: HISTORICAL REVIEW

### Yesterday

Thirty years ago, the discipline of sociology had a crisp concept of *profession*. It deemed a profession, in contrast to other occupations, to be a vocation with a body of knowledge and skills put into service for the good of others.<sup>2</sup> The specialized, complex, and uncertain nature of that expertise conferred autonomy on the profession charged with self-regulation to honor the social contract. Medicine was the profession *par excellence*.

Thirty years ago in medical education, the concepts of profession and professionalism, per se, were absent. Of course, there was interest in behaviors now labeled professional. But these behaviors were often treated as a residual category referring to anything that was “not cognitive.” Work on noncognitive characteristics of medical school applicants, medical students, and graduates illustrates this approach.<sup>3-6</sup>

In the early 1980s a major change occurred. The American Board of Internal Medicine (ABIM) began its humanism project.<sup>7</sup> It saw humanism as an entity consisting of respect, compassion, and integrity. It supported a number of studies for evaluating the humanism of residents. In turn, the humanism initiative led to Project Professionalism in the mid-1990s.<sup>8</sup>

### Today

The concept of professionalism in medical education today is clearly circumscribed with specific elements. Definitions, empirically and prospectively derived, abound.

A variety of methods has yielded empirical definitions of professionalism. For example, a survey of over 1,500 respondents identified 87 positive and 29 negative physician qualities, many involving professionalism.<sup>9</sup> A critical-incident technique implicated five noncognitive skills in professionalism.<sup>10</sup> A process of normative consensus specified 13 traits related to professionalism.<sup>11</sup>

Prospectively, about 50% of medical schools have written criteria and specific assessment methods to assess professional behavior.<sup>5,12</sup> For example, the professional development assessment form at Northeastern Ohio Universities College of Medicine lists eight elements of professionalism, along with definitions and specific questions for assessing medical students.<sup>13</sup> The elements are reliability and responsibility, honesty and integrity, maturity, respect for others, critique, altruism, interpersonal skills,

and [absence of] impairment (i.e., psychological/chemical). The student evaluation form at the University of New Mexico School of Medicine contains similar elements, in particular, reliability and responsibility, honesty and integrity, maturity, critique, and impairment; but it includes communication skills and respect for patients while omitting altruism.<sup>14</sup> The physicianship evaluation form at the University of California, San Francisco School of Medicine, directs evaluators to gauge just four aspects of professionalism: professional responsibility, self-improvement and adaptability, relationships with patients and families, and relationships with members of the health care team.<sup>15</sup>

Professional organizations, such as the American Board of Internal Medicine,<sup>8</sup> the Society of Academic Emergency Medicine,<sup>16</sup> and the Accreditation Council on Graduate Medical Education,<sup>17</sup> have also defined professionalism prospectively, and they largely agree on the elements that compose it. These elements (both reminiscent of and distinct from those that appear on school forms cited above) are altruism; respect for other people; additional humanistic qualities; honor, integrity, ethical and moral standards; accountability; excellence; and duty/advocacy.<sup>8,16,17</sup> According to the ABIM,<sup>8</sup> altruism demands that the best interests of patients, not self-interest, guide physicians. Respect for others (ranging from patients to medical students) is the essence of humanism. Honor and integrity entail the highest standards of behavior and the refusal to violate one's personal and professional codes. Accountability, at multiple levels, includes fulfilling the contract governing the doctor-patient relationship, the profession, and society. Excellence entails a commitment to exceed ordinary expectations and commitment to lifelong learning. Duty is the free acceptance of commitment to service.

Leaders in medical education concur

with these elements and their definitions, but only up to a point. Several of them include autonomy and self-regulation<sup>18–21</sup> as well as uncertainty in the mix of elements.<sup>22</sup> Authors and organizations also vary in the emphasis they give to some of the elements. Altruism is the lynchpin for the ABIM.<sup>8</sup> Duty, advocacy, service, and social responsiveness are central to the perspectives of a number of authors.<sup>18–24</sup> Although autonomy and self-regulation may be passé due to the encroaching role of agencies external to medicine, several authors strongly contend that these elements are more critical than ever if medicine is to remain a profession.<sup>18–21</sup> Humanism should be treated as an entity, whose central concept is empathy.<sup>25</sup>

There are nuanced differences as well. Interpretations of accountability vary across authors.<sup>8,26,27</sup> Overlaps between elements exist. Humanism, for example, includes integrity in the ABIM schema<sup>7</sup>; yet integrity stands as a separate entity in its later formulation of professionalism.<sup>8</sup> The AAMC's Medical Schools Objective Project (MSOP) does not contain the concept of professionalism, but it speaks clearly to its elements.<sup>28</sup>

Additionally, challenges to professionalism have been recognized. Conflicts of interest, abuse of power, lack of conscientiousness, destructive arrogance—these and other challenges are important for the assessment of professionalism.<sup>8,29–32</sup>

### Implications for Tomorrow

Medical education is no longer silent about the concept of professionalism. The literature offers core definitions that can serve as the foundation for next steps in research studies and in the development of assessment tools. Nuanced differences need clarification. Ongoing review of the definition of professionalism, such as that undertaken by the AAMC's Group on Educational Affairs,<sup>33</sup> the AAMC's Council of Aca-

demic Societies,<sup>34</sup> and the physician charter project of the American Board of Internal Medicine Foundation/American College of Physicians–American Society of Internal Medicine Foundation/European Federation of Internal Medicine<sup>35</sup> will be necessary to assure that the concept will be appropriate to the evolving needs of the world's populations for health care and well-being.

### MEASUREMENT OF PROFESSIONALISM

No single method exists for the reliable and valid evaluation of professional behavior. There are at least three types of studies, however, that may point the way for future evaluation thrusts. Some work evaluates professional behavior as part of clinical performance. Other studies evaluate only professional behavior, as a comprehensive entity in and of itself. Still other research evaluates single elements of professional behavior such as humanism, self-assessment, dutifulness, altruism, empathy and compassion, honesty, integrity, and ethical behavior, as well as communication. I describe measurement tools from each of these types of studies, report psychometric properties and substantive findings, and draw implications for next steps.

#### Measurement of Professionalism in Research on Evaluating Clinical Performance

Of interest here are studies of medical students and residents evaluating their peers and studies of practicing physicians evaluating their peers, residents, and medical students.

**Learners' assessments of peers.** Peer evaluation among medical students and residents may be an excellent source of information about the professional and nonprofessional behaviors of learners, for peers are in frequent, close contact with each other when no-one in authority is present. Although a notable

study describes a successful annual nomination of top peers in a medical school graduating class,<sup>36</sup> peer assessment typically depends upon rating scales. Internal consistency of these rating scales can be high.<sup>37</sup> Inter-rater reliability is moderate.<sup>38</sup> However, peer assessments may be subject to a halo effect,<sup>37</sup> since learners might not differentiate between peers' technical knowledge and skills and peers' professional behaviors. The relationship between peer assessments and faculty measures is weak to moderate,<sup>37,38</sup> although without a "gold standard" of professional behavior the import of this finding is unclear. Then too, peers may be reluctant to assess each other.<sup>39–41</sup>

On the other hand, peers do offer solid information about each others' interpersonal skills.<sup>39,42,43</sup> They contribute unique insight into the professional behavior of each other.<sup>36,44</sup> Through peer assessment, regulation of the medical profession from within by its own members may be achieved.

*Implications for tomorrow.* Peer assessment of professional behaviors holds promise. To be most useful for our purposes, peer-assessment tools should not include all the dimensions of clinical performance; rather their scope should be limited to professional behaviors only, due to the aforementioned halo effect. Psychometric properties of these tools need improvement. To do that, we need to understand peers' reluctance to evaluate each other. Such understanding can come from exploring peers' ideas about conditions conducive to their participation in peer assessment.

**Physicians' assessments of colleagues, residents, and medical students.** Studies of physicians evaluating professional behavior as a part of clinical performance have also relied largely on rating scales.<sup>45–48</sup> For example, an excellent study of physicians in practice assessing their peers<sup>45</sup> used a form that contained rating scales with items about knowledge, clinical skills, management of problems, and problem solving, on

---

the one hand, and on the other, respect, compassion, responsibility, and psychosocial aspects of care.

Generally, inter-rater reliability is poor in these studies, in part a reflection of the small numbers of raters used. To achieve an acceptable level of reliability, one study found that ratings from at least 11 physician associates of each physician subject would be needed.<sup>45</sup> Inter-rater agreement on humanistic items can be particularly low.<sup>49</sup>

Moreover, high intercorrelations across categories of behaviors often occur.<sup>50-52</sup> At best, raters make a distinction between technical knowledge and skills and professional/humanistic behaviors, according to a series of studies that consistently found a two-factor structure in clinical-performance rating data.<sup>45,53-56</sup> This result suggests that expert evaluators may cognitively separate their perceptions of learners and peers into just two categories without making distinct judgments among the separate elements of professional behavior,<sup>57</sup> although occasionally three factors have been derived from clinical-performance data.<sup>58,59</sup>

*Implications for tomorrow.* This prospect raises some quandaries for development of future measurement tools. If expert evaluators, in fact, organize their perceptions into a technical knowledge/skill category and a professional-behavior category, then we may not be able to obtain ratings of each of the various elements of professionalism in order to certify professional behaviors across the continuum of medical education. For certification, perhaps only summative judgments that a physician is or is not professional will suffice. However, to guide growth in professionalism along the continuum of medical education, we may profit from judgments of separate elements of professionalism to pass along to learners as formative evaluation. For that reason I would not choose a tool measuring clinical performance to assess professional behavior.

### Measurement of Professionalism as a Comprehensive Entity

Some studies exclusively focus on measuring professional behavior, in and of itself, by using a comprehensive definition of professionalism. This work falls into two categories. One type assesses groups of learners through surveys. The other evaluates individuals through critical-incident techniques.

**Surveys to measure professionalism of groups.** An outstanding study of the professional behaviors of groups of learners tackles forthrightly the issue of whether professionalism can be measured in medical education.<sup>60</sup> Students and residents from five institutions responded to questionnaire items that described professional and unprofessional behaviors of residents. The items, 12 in all, operationalized each of the ABIM's six elements of professionalism.

The internal reliability of the instrument was acceptable overall ( $\alpha = .71$ ). Factor analysis of the data yielded not a two-factor but a three-factor solution. Together these factors—labeled excellence, honor/integrity, and altruism/respect—explained 51% of the variance. Only the first factor, excellence, had acceptable internal reliability ( $\alpha = .72$ ); and it distinguished levels of excellence among residents in the five participating institutions. The remaining two and less reliable factors had too few items as well as items that overlapped. Further, an important altruism item loaded on the excellence factor rather than on the altruism/respect factor.

*Implications for tomorrow.* The findings of the study just reviewed suggest that respondents can distinguish among the elements of professionalism if the tool examines only professional behaviors. Although the study produced group scores that learners who might not yet have been expert raters assigned, its results are encouraging. Further work along this line should try to increase reliability of the instrument

across raters, ratees, and time. In doing so, it might well verify whether the items reflect not only learners' ideas about their everyday experiences of professional behaviors but also their ideas about the elements of such behaviors. Continued re-evaluation of a more general professionalism scale that also produced distinct dimensions of professionalism<sup>61</sup> might be productively applied to the medical education setting.

**Critical-incident techniques to measure professionalism of individuals.** The second line of inquiry in studies assessing professional behavior with a comprehensive definition entails the use of critical incidents to characterize individual learners. In a distinguished series of studies, Rhoton<sup>62-64</sup> qualitatively analyzed faculty narratives and comment cards for critical incidents of residents' behaviors. She transformed her qualitative categories into z-scores for subsequent quantitative analysis. Thereby she identified residents with unprofessional behaviors, although it is important to note that the faculty rarely labeled these residents as below par. She also described types of unprofessional behaviors. The most frequent types entailed expressions of personality problems, fabrication, and abdication of responsibility. She obtained predictors of unprofessional behaviors. These included deficiencies in conscientiousness, taking instructions, eagerness to learn, and efficiency. Finally, she found that residents with no instance of unprofessional behavior in their records achieved excellent clinical performances. But those with unprofessional behaviors performed poorly. Studies using other methods such as rating scales to measure professionalism<sup>65-67</sup> have found similar relationships between professional behavior and overall performance.

Additional work using critical incidents to assess professional behavior has entailed longitudinal assessment that tracks students' professionalism throughout their medical school stays with

the goal of remediation, if necessary.<sup>68-70</sup> The approach enables faculty to quantify their impressions of problematic students in a uniform manner on a form listing behavioral indicators of traits. The faculty form, reporting a student's unprofessional behavior in or outside class, goes to a dean who meets with the student and decides on appropriate action. Students received citations most often for lack of conscientiousness and poor relationships with the health care team. Over four years, reports were forwarded to a dean for 1% of students in one school and 2% in another school.<sup>69,70</sup>

In these two schools, the evaluation process itself provides for a measure of reliability, since at least two reports must reach a dean before she or he takes action and meets with the identified student for further exploration of the incident. Validity data come from case disposition. In one school, the dean found cause to take action in nine of the ten cases. In the other school, the dean took action in all five instances.

These longitudinal assessments highlight the problem of quantifying professional and unprofessional behaviors. Some behaviors are not quantifiable along a scale. Can a learner be just a little bit honest or score a "five" in integrity? Quantification is difficult, too, because unprofessional behaviors do not happen frequently. One of the programs just discussed found a potential way around the difficulty by using negative anchor points along a *severity* scale. In both programs the dean addressed the *significance* of the unprofessional behavior.

Other issues with these types of programs include their having a focus on unprofessional behaviors only. Accordingly, all students do not receive feedback. The absence of a report about a student's behavior is not necessarily a testament to that student's professional conduct. Faculty may be wary of the longitudinal assessment program, but they do participate.

*Implications for tomorrow.* Further investigations of assessing professional behaviors through critical incidents should proceed. Past studies using critical incidents point up the important role of the dispassionate, disciplined reviewer of behavior—be that person a researcher or a dean. They demonstrate the rich contribution that qualitative analysis can make to assessing unprofessional behaviors. If reports of incidents were to include the less severe and less significant along with major lapses in professional behavior over time, patterns in behavior could be identified. The usefulness of critical-incident techniques might also be expanded if reports about professional behaviors were also sought.

#### Measurement of Separate Elements of Professionalism

Of interest here is research on specific elements such as humanism; self-assessment, self-regulation, and self-reflection; as well as altruism, duty, empathy, and ethical decision making.

**Humanism.** Humanism has been evaluated through self-reports, objective structured clinical examinations (OSCEs), and rating scales. Several questionnaires eliciting self-reports have been developed to characterize humanistic trends among groups of learners.<sup>71,72</sup> The questionnaire of Abbott is noteworthy. Based upon Pellegrino's concept of humanism, its psychometric properties have been thoroughly established.

An OSCE has contributed to ascertaining whether the humanism of family medicine clerks can be predicted.<sup>73</sup> Standardized patients used an eight-item checklist derived from a recognized scale<sup>74</sup> to score clerks' humanism. The psychometric properties of the OSCE station were acceptable. Students' humanism scores bore a relationship to their scores on a reliable and valid measure of the values they placed on biopsychosocial aspects of care, early

in medical school and before the clerkship began. Communications OSCEs also come close to measuring humanism through checklists with items such as "greet you warmly."<sup>75</sup>

However, much of what we know about measuring professionalism stems from studies of humanism,<sup>76-81</sup> sparked especially by the ABIM in the 1980s. Faculty used either one global item or an array of items, each addressing a component of humanism—integrity, compassion, and respect—to rate residents. In turn, these ratings were compared with nurses' and patients' ratings of residents' humanism. Such ratings are unreliable unless large numbers of raters are used.<sup>82</sup> For example, if faculty used one global item, about 50 observations would be required to achieve an acceptable level of reliability. To obtain reliable ratings from nurses, between ten and 20 observations would be necessary; to obtain reliable ratings from patients, 50 observations would be needed.

Further, the humanism ratings that faculty gave to residents most often had little, if anything, to do with the humanism ratings nurses and patients gave to residents. The strongest relationship reported, of .7, was found between faculty and nurse ratings in just one study.<sup>83</sup> Humanism of a resident, it seems, depends on whom you ask.

A number of factors may account for the discrepancies between ratings, in addition to the small numbers of raters used in many of these studies. These factors include differential opportunities for observation. For example, in outpatient settings the difference between patients' and faculty members' ratings of residents' humanism was diminished.<sup>84</sup> Furthermore, different raters used different criteria<sup>85,86</sup>; for example, faculty stressed technical criteria, while patients made no distinction between technical competence and humanism. Then too, patients and nurses responded to instruments different from those the faculty used.<sup>82</sup> Moreover, hu-

manism scores given to residents varied by the gender of raters and the gender of ratees.<sup>82</sup> Women patients thought the care of men residents was more humanistic, for example; men patients thought more highly of the care of women residents. Men faculty held women residents to higher standards.<sup>87</sup> Finally, humanism scores also depended upon the ethnicity of raters<sup>88</sup> and the age and health status of patients.<sup>82</sup> Older, less sick patients viewed residents' humanism more positively.

*Implications for tomorrow.* These studies dramatically dismiss the notion that measuring humanism, indeed professionalism, is simple. To achieve reliable and valid ratings, considerable effort will be required. They show that no single perspective about the humanism of a physician may be adequate. They prompt the recommendation that a profile on humanism containing information about the physician or learner from multiple sources may be necessary and useful.

**Self-assessment, self-regulation, and self-reflection.** Self-assessment of professional behavior may be suspect.<sup>89</sup> Indeed, self-assessment of technical knowledge and skills may not be accurate.<sup>90,91</sup> The relationship between self-assessments of relatively weak students and instructors' ratings of their performances is not as strong as the association between self-assessments of the better students and their instructors' ratings.<sup>92</sup> Residents' self-ratings of humanism are related weakly to others' ratings of their humanism, if at all.<sup>93,94</sup> Men and women self-assessed differently such personal attributes as caring and compassion, tolerance of uncertainty, and ability to work in a team,<sup>95,96</sup> although this finding is not consistent across investigations.<sup>97</sup> A new relative-ranking technique appeared promising in the self-assessment of interviewing skills.<sup>98</sup> However, when residents, using the relative-ranking technique, self-assessed a broad range of their clinical performances in a patient-care setting,

they said they most needed to increase their knowledge and skills but saw relatively little reason to improve their collegiality and team relationships.<sup>99</sup> Further, learners are reluctant to rate themselves.<sup>90,91</sup> The bias of social desirability is strong in measuring professionalism, and it may be rampant in self-assessment.

Self-assessment, however, can be accurate under certain conditions<sup>90,91</sup>; namely, when faculty expect learners to gather and interpret data on their performances and when they formally require students to reconcile self-assessments with credible external evaluative sources.

*Implications for tomorrow.* Although self-assessment of professional behaviors may be difficult, work on measuring this skill with regard to professionalism needs to continue. Self-assessment is a critical component of professionalism. The relative-ranking technique deserves further study. As with measuring other elements of professionalism, identification of the conditions that could support accurate self-assessment is vital.

**Other specific elements of professionalism.** Tools are available to provide a basis for assessing such elements of professionalism as altruism,<sup>100-102</sup> duty<sup>103</sup> and service,<sup>104,105</sup> empathy,<sup>106-109</sup> and ethical decision making.<sup>110,111</sup> Personality and value inventories<sup>103,112-115</sup> or tests of moral reasoning<sup>110,111</sup> with excellent psychometric properties exist. But some of them might not be directly germane to medical education. Research on empathy training<sup>116</sup> and studies of ethical dilemmas<sup>117</sup> point in that direction.

On the other hand, OSCEs that test learners' ethical reasoning, ethical behaviors, and communication skills might have greater clinical relevance. Studies have found that students' performances in communication increased through time.<sup>118</sup> A low rating from a standardized patient in a communications OSCE is rarely related to a high rating from a real patient in the clinical

setting.<sup>119</sup> OSCEs also can test the ability to convey empathy.<sup>120</sup>

Yet, OSCEs have been criticized for artificiality.<sup>121</sup> Ethical decision making can be inextricably entwined with communication skills.<sup>121</sup> Further, any single station has low reliability.<sup>122,123</sup> Scores are confounded by the content of the stations.<sup>122,124</sup> A recent exploratory study of a three-station OSCE in which standardized patients rated medical students' communication skills and professionalism along with their technical performance also found case specificity.<sup>125</sup> Further, the study obtained little relationship between the standardized patients' professionalism ratings (based on three items) and other potential indicators of professional behavior.<sup>125</sup>

*Implications for tomorrow.* Standard psychological tests with outstanding psychometrics may be an excellent resource for measuring altruism, duty, empathy, and ethical and moral reasoning. The role they could play in future assessment of professional behavior should be explored. Their potential may be maximized if they are framed to reflect the clinical setting.

Standardized patients in OSCE settings can establish learners' competence in ethical reasoning, ethical behavior, and communication. Since OSCEs with standardized patients can mimic clinical situations, their usefulness in assessing other elements of professional behavior should be studied. Because their reliability depends on the number of rating opportunities, however, the effort needed to generate solid tests of the elements of professional behavior by using OSCEs will be considerable.

## RECOMMENDATIONS TO IMPROVE FUTURE ASSESSMENTS OF PROFESSIONALISM

### Professional Behaviors as Value Clashes

More than one author enjoins us to stress behaviors in assessing profession-

alism.<sup>126,127</sup> In searching for ways to improve assessment of professionalism, the innovative review of Ginsburg and colleagues<sup>127</sup> notes that traditional evaluation methods rely on abstract idealized definitions that characterize people, rather than their behaviors, as unprofessional or professional and imply that professionalism is stable. Yet several studies question the stability of professionalism.<sup>128–130</sup> For example, through administration of the Minnesota Multiphasic Personality Inventory, serious personality disorders were discovered in two psychiatry residents who subsequently lost their licenses for professional misconduct. Other residents also showed the same personality traits; yet no reports were lodged against them in 15 years of follow-up.<sup>130</sup>

Ginsburg and colleagues<sup>127</sup> believe that measures of stable professional traits also miss the mark because they do not view unprofessional behaviors as expressing clashes between two or more equally worthy values. Evidence can be found for the concept that value conflicts underlie unprofessional behaviors. For example, in discussing ethical dilemmas with peers, medical students struggled with several conflicts between worthy values that led to questionable behaviors.<sup>131–133</sup> These dilemmas included conflicts between learning medicine by practicing on patients and providing care to patients, between honesty and integrity and being a good team player, and between talking with patients to gain social knowledge and talking with patients to gain medical knowledge to become a competent physician.<sup>131</sup> During a Group on Educational Affairs discussion group, Teaching, Learning, and Assessing Altruism and Dutifulness (part of the 1999 AAMC annual meeting in Washington, D.C.), faculty observed that among students, residents, and their own colleagues the values of conscientiousness and excellence could easily conflict with altruism. A survey revealed value clashes between care and ethics, on the

one hand, and money, on the other, that practicing physicians encounter from participating in two potentially opposed social structures—medicine and managed care.<sup>134</sup> The survey also described some resulting, and less-than-model, behaviors.<sup>134</sup>

Theories also suggest that the thoughtful seasoned practitioner must come to grips with moral ambiguity, value conflicts, and ethical dilemmas.<sup>135</sup>

Building upon the notion that professional behavior is an expression of value conflict is research describing OSCEs that require students to respond to difficult communication tasks.<sup>136</sup> The success of these OSCEs led to the suggestion that they could place students in situations involving difficult value conflicts where their responses might reveal professional lapses.<sup>127</sup>

Ginsburg and colleagues<sup>127</sup> also maintain that how learners resolve the conflict between values is as important as the behavior itself. Their suggestion is reminiscent of several attempts to evaluate professional behavior already in the literature. These include a written follow-up to an ethics OSCE station where students explained their choices of actions<sup>137</sup> and a professional decisions inventory<sup>138</sup> in which students indicated how they would respond in a clinical scenario and then chose values to justify their responses. Learners' think-aloud exercises, narrative reports,<sup>139,140</sup> responses to cases,<sup>141</sup> reflective pieces,<sup>142</sup> and focus-group transcripts can be subjected to qualitative analysis to lay bare resolution of value conflicts.

Examining this process of resolution is a critical step. Not only can it divulge how a learner deals with the conflict, it also can disclose whether the learner perceives a conflict in the first place and how and why the values the learner uses might deviate from the elements of professional behavior. Illustrative of this point is the sharp division that occurred between students and faculty after discipline was imposed upon students who

perceived they had done nothing wrong when one offered to write a paper in a health policy course for another.<sup>143</sup> The distance between generations and the diversity of our students in this post-modernist world further underscore the need for exploring learners' resolutions of value conflicts. Moreover, such exploration could reveal that some unprofessional behaviors might not reflect value clashes but instead reflect other etiologies.

*Implications for tomorrow.* The hypothesis should be tested that measurement tools focused on professional behaviors as expressions of value conflicts will produce more reliable and valid instruments. Such tools might be especially useful for evaluating "routine" occasional lapses into unprofessional behaviors. Research on the process of resolving value conflicts should also be continued. The efficacy of techniques such as qualitative analysis of reflective pieces should be investigated. The technique of moral conversation, where participants strive to see the worth in others' arguments and the flaws in their own, might also provide insight into value clashes that our learners in medical education face.<sup>144</sup>

### **Professional Behaviors as Context-dependent: Stages of a Medical Career**

Ginsburg et al.<sup>145</sup> also argue that assessment in the area of professionalism must recognize the specificity of professional behaviors. Much evidence for their proposition—that professional behaviors depend on context—can be found in studies of ethical dilemmas,<sup>146</sup> where values that the same individuals brought to bear in taking actions varied across the scenarios presented. Research on peer assessment suggests situational specificity of unprofessional behaviors, since the frequency of peers' reports of their colleagues' negative behaviors was related to the quality of leadership on the health care team.<sup>147</sup> More specifi-

---

cally, groups with leaders who were physically absent or who used laissez-faire techniques had a greater frequency of negative peer reports than did other teams with leaders who were present and who unambiguously communicated their expectations for group members.

The clearest suggestion in the literature that professional behavior may be context-dependent comes from studies of stages or phases of medical careers. According to a study of the dreams of medical students and residents, critical episodes during training produced psychological defenses that regularly reduced and then increased learners' abilities to interact with patients empathically and altruistically.<sup>148</sup> Expressions of empathy and regard among residents in a support group waxed and waned during the first year, with a rise in empathy noted during the most stressful months of the year when professional problems were more frequently discussed.<sup>149</sup> Cross-sectional, longitudinal, and retrospective studies of cynicism and humanism illustrate similar ups and downs.<sup>150-153</sup> Students felt they grew more cynical during medical school but also more interested in and helpful with patients.<sup>153</sup> Medical students were found to be most cynical, while residents and especially faculty were less so.<sup>150</sup>

From these results flows the notion that what physicians need to learn and thus what needs to be assessed regarding professional behavior will vary according to career stage. Medical educators at the AAMC<sup>154</sup> selected from all the elements of professional behavior the following as most applicable to medical students: altruism, ethical and moral standards, responsiveness to society's needs, and core humanistic values. On the other hand, educators who subscribe to the effectiveness of anticipatory socialization would argue that all of the elements of professional behavior should apply to medical students, rather than just those most relevant to the student role. For residents, the Accredita-

tion Council for Graduate Medical Education has specified the elements of professionalism that were presented at the outset of the present article. Yet residents themselves have defined professionalism as entailing competence, first and foremost.<sup>155</sup> In their view of professionalism they accorded much less importance to values such as altruism and social responsiveness.<sup>155</sup> Teachers of residents have taken a similar stance.<sup>156</sup> Whether students and residents should be assessed along all of the elements or only those that more directly bear on their roles is a critical next step in assessment of professional behaviors.

Additionally, the level of learning that will be expected of medical students and residents is an unresolved but important issue. Useful here is Miller's pyramid model of learning<sup>157</sup> that suggests corresponding levels of assessment: knowledge, capacity to apply, and actualization in practice—the “know, can, do” schema. Many of the objectives in the MSOP report are cast only in terms of knowing or understanding. A study showing that medical students' experiences with professional behavior do not correspond well with standard abstract definitions highlights the need for instruction in knowledge about the elements of professionalism.<sup>158</sup> Several approaches to the longitudinal assessment of the ethical development of medical students<sup>159</sup> and residents<sup>160</sup> provide for testing of knowledge as well as competence and actualization in practice.

*Implications for tomorrow.* Two issues await resolution through future research. First, should all or only some elements of professionalism be assessed at different stages of a medical career? For example, should medical students be held accountable for only those elements of professionalism, such as responsibility, that are most germane to their learning role? Second, should all levels or only some levels of assessment be used during the various stages of a

medical career? Should medical students be held responsible for knowledge about professionalism only, residents for knowledge and competence, and practicing physicians for knowledge, competence, and application? What are the relationships between a learner's knowledge of professionalism, demonstrated competence, and application? Under what conditions does it follow that if physicians demonstrate knowledge and competence in professionalism, they will act professionally?

Perhaps a matrix should be developed to indicate which levels of assessment will be applied to which elements at which career stage. Perhaps only indicators of each element will vary by stage of career. Yet the literature enjoins an emphasis on the third level of “doing” because of the issue of social desirability that attends the assessment of professional behaviors. In tests of knowledge, on OSCEs, in essays, and in entries in diaries and journals, learners may display competence in professionalism. But when confronted in the heat of the moment with value conflicts, they may lapse into unprofessional behaviors. Actions speak louder than words. However, the “know and can” levels of the pyramid should not be neglected.

### **The Environment in Which Assessment Occurs**

Pursuing the above recommendations will not be sufficient, however, to generate measurement tools that produce dependable, credible, and transferable information about learners' professionalism unless we also consider the environment in which the assessment of professionalism occurs. We need to pay attention to the institutional stance on assessing professional behavior and to the conditions under which the assessment is administered if we are to be successful in our quest for improved assessment of learners' professionalism.

**The institutional stance.** Theoretical and empirical studies rivet our attention



on the hidden and informal curriculum.<sup>161-164</sup> A great deal of teaching about professional values occurs outside scheduled class time, in the informal curriculum, when faculty are absent.<sup>165</sup> Further, only some of the “teaching” there is congruent with the announced professional values of an institution. According to one study, the taught, or informal, curriculum spoke to the *burden* of service and interprofessional *disrespect* rather than to their opposites (the *importance* of service and interprofessional *respect*) that the recommended, or formal, curriculum espoused.<sup>166</sup> Furthermore, the taught curriculum emphasized industriousness of learners, while the recommended curriculum was silent on that matter.<sup>166</sup> Only if the lessons of the informal curriculum are clearly understood can we incorporate authentic indicators of professional and unprofessional behaviors into our measurement tools.

Moreover, the reticence of students, residents, faculty, and colleagues to report unprofessional behaviors must be addressed.<sup>167-171</sup> To do that, we need to know the stance of the informal culture on assessing professional behavior. Just how important is such assessment? Is it considered an inconvenience, a necessary evil, or a vital link in enabling all members of our academic health centers to become and be professional? Is there courage to follow through with discipline, if necessary? Do all levels of administration embrace the assessment of professional behavior or merely pay lip service to it?

### Administration of Assessment

In light of the reluctance to assess professional behavior, our efforts will also need to explore the conditions that will encourage participants to provide insightful, credible, dependable information. Some of these circumstances include the spirit in which assessment proceeds. Is it carried out against a

backdrop of primary prevention and health promotion or against one of transgression, sickness, and deviance? Is it done in the spirit of justice and the social good or only for the good of the individual?

Other conditions may be important as well. The purpose of the assessment comes to mind. Is the purpose formative, for guidance, growth, and striving toward the ideal? Or is it summative? Is the purpose clearly communicated to learners and faculty? Is the assessment tool compatible with the purpose of the assessment?

Then there is the breadth of the assessment. Is it directed toward identifying both unprofessional and professional behaviors? Are patterns of behaviors or just one-time lapses of interest? How does the assessment take into account the severity of the lapses? How much unprofessional behavior is too much? Does the assessment attempt to uncover the reasons for the behavior as well as the behavior itself? Is all unprofessional behavior a matter of clashes between worthy values? Do learners recognize their behaviors as professional or unprofessional?

The types of people involved in the assessment may be yet another consideration. Who does the evaluation? Do the most vulnerable people in the system, such as patients or ward secretaries, have input? How do the assessors' attitudes and perceptions of professionalism affect the quality of information they report about learners? Who receives the evaluation—a credible fair reviewer? Does the assessment entail an individual or group decision? Faculty in one department were more likely to identify lapses in professional behavior when they discussed learners in a committee meeting than they were on checklists and in a comment section of an evaluation form.<sup>172</sup>

The safety of participants in assessing professional behavior is another key aspect of the environment. What constitutes safety for learners and faculty? Is

the assessment anonymous, confidential, or signed?

Finally, do all participants receive education in the assessment of professional behavior?

Noting the dissatisfaction with evaluation systems in residencies, Gordon<sup>173</sup> offers a proposal that splits the evaluation process in two. The proposal may be worth considering in the context of professional behavior. One system, for monitoring standards to assure that learners do not fall below established standards, is the faculty's responsibility. The other, for professional growth and development beyond the minimum, is the responsibility of the residents. The proposal assumes that both faculty and residents are legitimate decision makers concerning a resident's education. The quality control system of the faculty would use simple qualitative measures to screen for residents' adherence to minimum standards, give early warning, and provide rapid follow-up. The resident-controlled, guidance-oriented system would concern itself with professional growth, self-assessment, reflection, and peer and faculty coaching. Faculty would insist only that residents participate in good faith.

*Implications for tomorrow.* Just as we need to study the context in determining whether behavior is professional or unprofessional, so too do we need to systematically examine how the environment influences the quality of the information our assessment tools yield about the professional behaviors of learners. The acceptability and efficacy of Gordon's proposal should be studied in the context of professional behaviors across the continuum.

### FINAL THOUGHTS

Throughout this article, I have outlined steps that could be taken to strengthen the assessment of professionalism in the future. For the reader's convenience, I list them below:

- The well-circumscribed concept of professionalism can serve as a foundation for future measurement initiatives, but it does require clarification.
- Assessment of professionalism should focus on professionalism, in and of itself.
- Instruments that measure the separate elements of professionalism should be developed.
- Rigorous qualitative approaches to assessment should be encouraged, along with more quantitative measures of the elements of professional behavior that might be derived from the use of standardized patients in OSCE settings, for example.
- The hypothesis should be explored that to improve assessment of professionalism, our tools should emphasize behaviors as expressions of value conflicts, explore the resolution of these conflicts, and take into account the contextual nature of professional behaviors.
- Of most immediate concern is whether measurement tools should be tailored to the stage of a medical career.
- How the environment can support or sabotage the assessment of professional behavior is also a central issue.

Without solid assessment tools, questions about the efficacy of approaches to educating learners about professional behavior will not be effectively answered.

An invitation to deliver the 2001 Jack L. Maatsch Lecture at Michigan State University College of Human Medicine, East Lansing, Michigan, and to explore the evaluation of professional behavior in the lecture was the impetus for this article. The author gratefully acknowledges the role of that invitation in writing the article. The author's participation in the project on professionalism of the Group on Educational Affairs of the Association of American Medical Colleges contributed to developing the paper. Finally, the discussions that she had with David T. Stern, MD, PhD, at the University of Michigan Medical School, helped to clarify the role of environment in assessment, especially regarding peer evaluation.

## REFERENCES

### INTRODUCTION

1. Stufflebeam DL, Foley WJ, Gephart WJ, et al. *Educational Evaluation and Decision-Making*. Itasca, IL: Peacock, 1971.

### THE CONCEPT OF PROFESSIONALISM: HISTORICAL REVIEW

2. Parsons T. *The Social System*. Glencoe, IL: The Free Press, 1951.
3. Calkins EV, Richards JM, McCauley A, Burgess M, Willoughby TL. Impact on admission to a school of medicine with an innovation in selection procedures. *Psychol Rep*. 1974;35:1135-42.
4. Keck J, Arnold L, Calkins EV, Willoughby TL. Efficacy of cognitive/noncognitive measures in predicting resident physician performance. *J Med Educ*. 1979;54:759-65.
5. Miller GD, Frank D, Franks R, Getto C. Noncognitive criteria for assessing students in North American medical schools. *Acad Med*. 1989;64:42-5.
6. Murden R, Galloway GM, Reid JC, Colwill JM. Academic and personal predictors of clinical success in medical school. *J Med Educ*. 1978;53:711-9.
7. Subcommittee on Evaluation of Humanistic Qualities in the Internist, American Board of Internal Medicine (ABIM). *Evaluation of humanistic qualities in the internist*. *Ann Intern Med*. 1983;99:720-4.
8. American Board of Internal Medicine (ABIM). *Project Professionalism*. Philadelphia, PA: American Board of Internal Medicine, 1994.
9. Price PB, Lewis EA, Loughmiller GC, Nelson DE, Murray SL, Taylor CW. Attributes of a good practicing physician. *J Med Educ*. 1971;46:229-37.
10. Rhoton MF, Barnes A, Flashburg M, Ronai A, Springman S. Influence of anesthesiology residents' noncognitive skills on the occurrence of critical incidents and the residents' overall clinical performances. *Acad Med*. 1991;66:359-61.
11. Benyamini K, Kedar HS, Raveh I. How do supervising doctors construe the medical student in clinical training? *Med Educ*. 1987;21:410-8.
12. Swick HM, Szenas P, Danoff D, Whitcomb ME. Teaching professionalism in undergraduate medical education. *JAMA*. 1999;282:830-2.
13. Gibson DD, Coldwell LL, Kiewit SF. Creating a culture of professionalism. *Acad Med*. 2000;75:509.
14. Phelan S, Obenshain SS, Galey WR. Evaluation of the non-cognitive professional traits of medical students. *Acad Med*. 1993;68:799-803.
15. Papadakis MA, Osborn EHS, Cooke M, Healy K. A strategy for the detection and evaluation of unprofessional behavior in medical students. *Acad Med*. 1999;74:980-90.
16. Adams J, Schmidt T, Sanders A, Larkin GL, Knopp R. Professionalism in emergency medicine. *Acad Emerg Med*. 1998;5:1193-9.
17. ACGME Outcome Project. *Enhancing Residency Education through Outcomes Assessment; General Competencies*. Version 1.3. (<http://www.acgme.org/outcome/comp/compFull.asp>). Accessed 2/22/02. Accreditation Council for Graduate Medical Education, Chicago, IL, 1999.
18. Cruess RL, Cruess SR. Teaching medicine as a profession in the service of healing. *Acad Med*. 1997;72:941-52.
19. Cruess RL, Cruess SR, Johnston SE. Renewing professionalism: an opportunity for medicine. *Acad Med*. 1999;74:878-84.
20. Cruess RL, Cruess SR, Johnston SE. Professionalism and medicine's social contract. *J Bone Joint Surg*. 2000;82A:1189-94.
21. Cruess RL, Cruess SR, Johnston SE. Professionalism: an ideal to be sustained. *Lancet*. 2000;356:156-9.
22. Swick HM. Toward a normative definition of medical professionalism. *Acad Med*. 2000;75:612-6.
23. Irvine D. The performance of doctors: the new professionalism. *Lancet*. 1999;353:1174-7.
24. Rothman DJ. Medical professionalism—focusing on the real issues. *N Engl J Med*. 2000;342:1284-6.
25. Hafferty FW. Keynote address. Overcoming the barriers to sustaining humanism in medicine: influencing the culture through a humanism honor society. Invitational Conference of the Arnold P. Gold Foundation, Secaucus, NJ, March 2001.
26. Emmanuel EJ, Emmanuel LL. What is accountability in health care? *Ann Intern Med*. 1996;124:229-39.
27. Evetts J. Professionalization and professionalism: issues for interprofessional care. *J Interprofessional Care*. 1999;13:119-28.
28. The Medical School Objectives Writing Group. *Learning objectives for medical student education—guidelines for medical schools: report I of the Medical School Objectives Project*. *Acad Med*. 1999;74:13-8.
29. Swick HM. Academic medicine must deal with the clash of business and professional values. *Acad Med*. 1998;73:751-5.
30. Kassirer JP. Medicine at center stage. *N Engl J Med*. 1993;328:1268-9.

31. Southwick F. Who was caring for Mary? *Ann Intern Med.* 1993;118:146–8.
32. Ingelfinger FJ. Arrogance. *N Engl J Med.* 1980;303:1507–11.
33. Arnold L, Bennett N, Distlehorst L, Smith C, Stern DT. Professionalism. Papers given at the annual meeting of the Association of American Medical Colleges, Washington, DC, November 2001.
34. Cooper TG. Professionalism: what it is and why does it matter? Paper given at the annual meeting of the Association for Behavioral Science and Medical Education, North Falmouth, MA, October 2001.
35. Blank L, Brennan T, McDonald W, Stobo J. Physician Charter on Professionalism: what and why? Papers given at the annual meeting of the Association of American Medical Colleges, Washington, DC, November 2001.

#### MEASUREMENT OF PROFESSIONALISM

##### Research on Evaluating Clinical Performance

###### *Learners' Assessment of Peers*

36. Small PA Jr, Stevens B, Duerson MC. Issues in medical education: basic problems and potential solutions. *Acad Med.* 1993;68(10 suppl):S89–S93.
37. Arnold L, Willoughby TL, Calkins EV, Gammon L, Eberhart G. Use of peer evaluation in the assessment of medical students. *J Med Educ.* 1981;56:35–42.
38. Panszi S, Gruppen L, Grum C, Stern DT. What do peers know about professionalism? Poster presentation given at the Research in Medical Education Conference, Group on Educational Affairs, AAMC Annual Meeting, Chicago, IL, 2000.
39. Helfer RE. Peer evaluation: its potential usefulness in medical education. *Br J Med Educ.* 1972;6:224–31.
40. Thomas PA, Gebo KA, Hellmann DB. A pilot study of peer review in residency training. *J Gen Intern Med.* 1999;14:551–4.
41. Van Rosendaal GMA, Jennett PA. Resistance to peer evaluation in an internal medicine residency. *Acad Med.* 1992;67:63.
42. Linn BS, Arostegui M, Zeppa R. Performance rating scale for peer and self-assessment. *Br J Med Educ.* 1975;9:98–101.
43. Schumacher CFJ. A factor analytic study of various criteria of medical examinations. *J Med Educ.* 1964;39:192–6.
44. Kubany AJ. Use of sociometric peer nominations in medical education research. *J Appl Psychol.* 1957;41:389–94.

###### *Physicians' Assessments of Colleagues, Residents, and Medical Students*

45. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA.* 1993;269:1655–60.
46. Leibbrandt TJ, Kukora JS, Dent TL. Integrating educational objectives and the evaluation process in a general surgery residency program. *Acad Med.* 2001;76:748–52.
47. Reisdorff EJ, Hayes OW, Carlson DJ, Walker GL. Assessing the new general competencies for resident education: a model from an emergency medicine program. *Acad Med.* 2001;76:753–7.
48. Willoughby TL, Gammon LC, Jonas HS. Correlates of clinical performance during medical school. *J Med Educ.* 1979;54:453–60.
49. Johnson D, Cujec B. Comparison of self, nurse, and physician assessment of residents rotating through an intensive care unit. *Crit Care Med.* 1998;26:1811–6.
50. Dawson-Saunders B, Paiva R. The validity of clerkship performance evaluations. *Med Educ.* 1986;20:240–5.
51. Davis JK, Inamdar S, Stone RK. Inter-rater agreement and predictive validity of faculty ratings of pediatric residents. *J Med Educ.* 1986;61:901–5.
52. Durand RP, Levine JH, Lichtenstein LS, Fleming GA, Ross GR. Teachers' perceptions concerning the relative values of personal and clinical characteristics and their influence on the assignment of students' clinical grades. *Med Educ.* 1988;22:335–41.
53. Ramsey PG, Carline JD, Blank LL, Wenrich MD. Feasibility of hospital-based use of peer ratings to evaluate the performances of practicing physicians. *Acad Med.* 1996;71:364–70.
54. Maxim BR, Dielman TE. Dimensionality, internal consistency, and inter-rater reliability of clinical performance ratings. *Med Educ.* 1987;21:130–7.
55. Gough HG, Hall WB, Harris RW. Evaluation of performance in medical training. *J Med Educ.* 1964;39:679–92.
56. Gertsma RJ, Chapman JE. The evaluation of medical students. *J Med Educ.* 1967;42:938–48.
57. Littlefield J, Arnold L, Barzansky B. Literature assessment: the validity of assessing physician professional behavior. Paper given at the fall meeting of the Society of Directors of Research in Medical Education, Washington, DC, November 1996.
58. Arnold L, Willoughby TL, Calkins EV. Towards understanding the clinical performance of physicians. *J Med Educ.* 1984;59:591–4.

59. Benyamini K, Kedar HS, Raveh I. How do supervising doctors construe the medical student in clinical training? *Med Educ.* 1987;21:410–8.

##### Measurement of Professionalism as a Comprehensive Entity

###### *Surveys to Measure Professionalism of Groups*

60. Arnold EL, Blank LL, Race KEH, Cipparone N. Can professionalism be measured? the development of a scale for use in the medical education environment. *Acad Med.* 1998;73:1119–21.
61. Kennedy KN, Ramsey R. Psychometric evaluation of the Hall professionalism scale. *Psychol Rep.* 1995;77:331–8.

###### *Critical-incident Techniques to Measure Professionalism of Individuals*

62. Rhoton MF. A new method to evaluate clinical performance and critical incidents in anesthesia: quantification of daily comments by teachers. *Med Educ.* 1989;23:280–9.
63. Rhoton MF, Barnes A, Flashburg M, Ronai A, Springman S. Influence of anesthesiology residents' noncognitive skills on the occurrence of critical incidents and the residents' overall clinical performances. *Acad Med.* 1991;66:359–61.
64. Rhoton MF. Professionalism and clinical excellence among anesthesiology residents. *Acad Med.* 1994;69:313–5.
65. Herman MW, Veloski JJ, Hojat M. Validity and importance of low ratings given medical graduates in noncognitive areas. *J Med Educ.* 1983;58:837–43.
66. Rowley BD, Baldwin DC Jr, Bay RC, Canula M. Can professional values be taught? A look at residency training. *Clin Orthop Rel Res.* 2000;378:110–4.
67. Sheehan TJ, Husted S, Candee D, Cook CD, Borgen M. Moral judgement as a predictor of clinical performance. *Eval Health Prof.* 1980;3:393–404.
68. Loeser H, Papadakis M. Promoting and assessing professionalism in the first two years of medical school. *Acad Med.* 2000;75:509–10.
69. Phelan S, Obenshain SS, Galey WR. Evaluation of the non-cognitive professional traits of medical students. *Acad Med.* 1993;68:799–803.
70. Papadakis MA, Osborn EHS, Cooke M, Healy K. A strategy for the detection and evaluation of unprofessional behavior in medical students. *Acad Med.* 1999;74:980–90.

## Measurement of Separate Elements of Professionalism

### Humanism

71. Abbott LC. A study of humanism in family physicians. *J Fam Pract.* 1983;16:1141-6.
72. Wolf TM, Balson PM, Faucett JM, Randall HM. A retrospective study of attitude change during medical education. *Med Educ.* 1989;23:19-23.
73. Rogers JC, Coutts L. Do students' attitudes during preclinical years predict their humanism as clerkship students? *Acad Med.* 2000;75(10 suppl):S74-S77.
74. Hauck FR, Zyzanski SJ, Alemagno SA, Medalie JH. Patient perceptions of humanism in physicians: effects on positive health behaviors. *Fam Med.* 1990;22:447-52.
75. Klamen DL, Williams RG. The effect of medical education on students' patient-satisfaction ratings. *Acad Med.* 1997;72:57-61.
76. Beckman H, Frankel R, Kihm J, Kulesza G, Geheb M. Measurement and improvement of humanistic skills in first-year trainees. *J Gen Intern Med.* 1990;5:42-5.
77. Blurton RR, Mazzaferri EL. Assessment of interpersonal skills and humanistic qualities in medical residents. *J Med Educ.* 1985;60:648-50.
78. Butterfield PS, Mazzaferri EL. A new rating form for use by nurses in assessing residents' humanistic behavior. *J Gen Intern Med.* 1991;6:155-61.
79. Linn LS, DiMatteo MR, Cope DW, Robbins A. Measuring physicians' humanistic attitudes, values, and behaviors. *Med Care.* 1987;25:504-13.
80. Linn LS, Cope DW, Robbins A. Socio-demographic and premedical school factors related to postgraduate physicians' humanistic performance. *West J Med.* 1987;147:99-103.
81. Merrill JM, Boisabuin EV Jr, Laux L, Lynch EC, Roessler R, Thornby JI. Measuring humanism in medical residents. *South Med J.* 1986;79:141-4.
82. Woolliscroft JO, Howell JD, Patel BP, Swanson DB. Resident-patient interactions: the humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors, and nurses. *Acad Med.* 1994;69:216-24.
83. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA.* 1993;269:1655-60.
84. McLeod PJ, Tamblyn R, Benaroya S, Snell L. Faculty ratings of resident humanism predict patient satisfaction ratings in ambulatory medical clinics. *J Gen Intern Med.* 1994;9:321-6.

85. Weaver MJ, Ow CL, Walker DJ, Degenhardt EF. A questionnaire for patients' evaluations of their physicians' humanistic behaviors. *J Gen Intern Med.* 1993;8:135-9.
86. Matthews DA, Feinstein AR. A new instrument for patients' ratings of physician performance in the hospital setting. *J Gen Intern Med.* 1989;4:14-22.
87. Klessig J, Robbins AS, Wieland D, Rubenstein L. Evaluating humanistic attributes of internal medicine residents. *J Gen Intern Med.* 1989;4:514-21.
88. Merrill JM, Boisabuin EV Jr, Cordova FA, et al. Culture as a determinant of humanistic traits in medical residents. *South Med J.* 1987;80:233-6.

### Self-assessment, Self-regulation, and Self-reflection

89. Ginsburg S, Regehr G, Hatala R, et al. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med.* 2000;75(10 suppl): S6-S11.
90. Gordon MJ. A review of the validity and accuracy of self-assessment in health professions training. *Acad Med.* 1991;66:762-9.
91. Gordon MJ. Self-assessment programs and their implications for health professions training. *Acad Med.* 1992;67:672-9.
92. Arnold L, Willoughby TL, Calkins EV. Self-evaluation in undergraduate medical education: a longitudinal perspective. *J Med Educ.* 1985;60:21-8.
93. Klessig J, Robbins AS, Wieland D, Rubenstein L. Evaluating humanistic attributes of internal medicine residents. *J Gen Intern Med.* 1989;4:514-21.
94. McLeod PJ, Tamblyn R, Benaroya S, Snell L. Faculty ratings of resident humanism predict patient satisfaction ratings in ambulatory medical clinics. *J Gen Intern Med.* 1994;9:321-6.
95. Clack GB, Head JO. Gender differences in medical graduates' assessment of their personal attributes. *Med Educ.* 1999;33:101-5.
96. Batenburg V, Smal JA, Lodder A, deMelker RA. Are professional attitudes related to gender and medical specialty? *Med Educ.* 1999;33:489-92.
97. Dornbush RL, Richman S, Singer P, Brownstein EJ. Medical school, psychosocial attitudes, and gender. *J Am Med Assoc.* 1991;46:150-2.
98. Regehr G, Hodges B, Tiberius R, Lofchy J. Measuring self-assessment skills: an innovative relative ranking model. *Acad Med.* 1996;71(10 suppl):S52-S54.
99. Harrington JP, Murnaghan JJ, Regehr G. Applying a relative ranking model to the self-assessment of extended performances. *Adv Health Sci Educ.* 1997;2:17-25.

### Other Specific Elements of Professionalism: Altruism, Duty, Empathy, Ethical Decision Making, and Communication Skills

100. Carlo G, Eisenberg N, Troyer D, Switzer G, Speer AL. The altruistic personality: in what contexts is it apparent? *J Pers Soc Psychol.* 1991;61:450-8.
101. Krebs D. Empathy and altruism. *J Pers Soc Psychol.* 1975;32:1134-46.
102. Sawyer J. The altruism scale: a measure of cooperative, individualistic, and competitive interpersonal orientation. *Am J Sociol.* 1966;71:407-16.
103. Magee M, Hojat M. Personality profiles of male and female positive role models in medicine. *Psychol Rep.* 1998;82:547-59.
104. Crandall SJ, Volk RJ, Loemker V. Medical students' attitudes toward providing care for the underserved. *JAMA.* 1993;269:2519-23.
105. Eckenfels EJ. Contemporary medical students' quest for self-fulfillment through community service. *Acad Med.* 1997;72:1043-50.
106. Davis MH. A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology.* 1980;10:85.
107. Evans BJ, Stanley RO, Burrows GD. Measuring medical students' empathy skills. *Br J Med Psychol.* 1993;66:121-33.
108. Latcher J, Mosek A. Similarities and differences between social work and medical students in empathy, conflict resolution, and professional image. *Social Sciences in Health, Int J Res Pract.* 1995;1:107-17.
109. Batenburg V, Smal JA, Lodder A, deMelker RA. Are professional attitudes related to gender and medical specialty? *Med Educ.* 1999;33:489-92.
110. Rest JR. *Development in Judging Moral Issues.* Minneapolis, MN: University of Minnesota Press, 1979.
111. Rest JR. *Moral Development: Advances in Research and Theory.* New York: Praeger, 1986.
112. Arnold L, Xu G, Epstein LC, Jones B. Professional and personal characteristics of graduates as outcomes of differences between combined baccalaureate-M.D. degree programs. *Acad Med.* 1996;71(10 suppl): S64-S66.
113. Epstein LC, Hayes J, Arnold L, et al. On becoming a physician: perspectives of students in combined baccalaureate-M.D. degree programs. *Teach Learn Med.* 1994;6:102-7.
114. Hojat M, Brigham TP, Gottheil E, Xu G, Glaser K, Veloski JJ. Medical students' personal values and their career choices a quarter-century later. *Psychol Rep.* 1998;83:243-8.

115. Eliason BC, Schubot DB. Personal values of exemplary family physicians: implications for professional satisfaction in family medicine. *J Fam Pract.* 1995;41:251-6.
116. Feighny KM, Arnold L, Monaco M, Munro JS, Earl B. In pursuit of empathy and its relation to physician communication skills: multidimensional empathy training for medical students. *Ann Behav Sci Med Educ.* 1998;5:13-21.
117. Rezler AG, Schwartz RL, Obenshain SS, Lambert P, Gibson JMcl, Bennahum DA. Assessment of ethical decisions and values. *Med Educ.* 1992;26:7-16.
118. Klamen DL, Williams RG. The effect of medical education on students' patient-satisfaction ratings. *Acad Med.* 1997;72:57-61.
119. Pieters HM, Touw-Otten FW, DeMelker RA. Simulated patients in assessing consultation skills of trainees in general practice vocational training: a validity study. *Med Educ.* 1994;38:226-33.
120. Schnabl GK, Hassard TH, Kopelow ML. The assessment of interpersonal skills using standardized patients. *Acad Med.* 1991;66(9 suppl):S34-S36.
121. Arnold RM, Forrow L. Assessing competence in clinical ethics: are we measuring the right behaviors? *J Gen Intern Med.* 1993;8:52-4.
122. Donnelly MB, Sloan D, Plymale M, Schwartz R. Assessment of residents' interpersonal skills by faculty proctors and standardized patients: a psychometric analysis. *Acad Med.* 2000;75(10 suppl):S93-S95.
123. Singer PA, Robb A, Cohen R, Norman G, Turnbull J. Performance-based assessment of clinical ethics using an objective structured clinical examination. *Acad Med.* 1996;71:495-8.
124. Hodges B, Turnbull J, Cohen R, Bienenstock A, Norman G. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ.* 1996;30:38-43.
125. Prislis MD, Lie D, Shapiro J, Boker J, Ra-decki S. Using standardized patients to assess medical students' professionalism. *Acad Med.* 2001;76(10 suppl):S90-S92.
126. Cohen JJ. Measuring professionalism: listening to our students. *Acad Med.* 1999;74:1010.
127. Ginsburg S, Regehr G, Hatala R, et al. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med.* 2000;75(10 suppl):S6-S11.
128. Sawyer J. The altruism scale: a measure of cooperative, individualistic, and competitive interpersonal orientation. *Am J Sociol.* 1966;71:407-16.
129. Carlo G, Eisenberg N, Troyer D, Switzer G, Speer AL. The altruistic personality: in what contexts is it apparent? *J Pers Soc Psychol.* 1991;61:450-8.
130. Garfinkel PE, Bagby RM, Waring EM, Dorian B. Boundary violations and personality traits among psychiatrists. *Can J Psychiatry.* 1997;42:758-63.
131. Christakis DA, Feudtner C. Ethics in a short white coat: the ethical dilemmas that medical students confront. *Acad Med.* 1993;68:249-54.
132. Feudtner C, Christakis DA, Christakis NA. Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Acad Med.* 1994;69:670-9.
133. Swenson SL, Rothstein JA. Navigating the wards: teaching medical students to use their moral compasses. *Acad Med.* 1996;71:591-4.
134. Castellani B, Wear D. Physician views on practicing professionalism in the corporate age. *Qualitative Health Research.* 2000;10:490-506.
135. Clark PG. Values in health care professional socialization: implications for geriatric education in interdisciplinary teamwork. *Gerontologist.* 1997;37:441-51.
136. Hodges B, Turnbull J, Cohen R, Bienenstock A, Norman G. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ.* 1996;30:38-43.
137. Smith SR, Balint JA, Krause KC, Moore-West M, Viles PH. Performance-based assessment of moral reasoning and ethical judgment among medical students. *Acad Med.* 1994;69:381-6.
138. Rezler AG, Schwartz RL, Obenshain SS, Lambert P, Gibson JMcl, Bennahum DA. Assessment of ethical decisions and values. *Med Educ.* 1992;26:7-16.
139. Branch WT. Deconstructing the white coat. *Ann Intern Med.* 1998;129:740-1.
140. Branch WT, Pels RJ, Lawrence RS, Arky R. Becoming a doctor: critical incident reports from third-year medical students. *N Engl J Med.* 1993;329:1130-2.
141. Branch WT, Pels RJ, Harper G, et al. A new educational approach for supporting the professional development of third-year medical students. *J Gen Intern Med.* 1995;10:691-4.
142. Lingard LA, Haber RJ. What do we mean by "relevance"? A clinical and rhetorical definition with implications for teaching and learning the case-presentation format. *Acad Med.* 1999;74(10 suppl):S124-S127.
143. Osborne E. Punishment: a story for medical education. *Acad Med.* 2000;75:241-4.
144. Nash RJ. *Answering the Virtuecrats: A Moral Conversation on Character Education.* New York: Teachers College Press, 1997.

### Professional Behavior as Context-dependent: Stages of a Medical Career

145. Ginsburg S, Regehr G, Hatala R, et al. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med.* 2000;75(10 suppl): S6-S11.
146. Rezler AG, Schwartz RL, Obenshain SS, Lambert P, Gibson JMcl, Bennahum DA. Assessment of ethical decisions and values. *Med Educ.* 1992;26:7-16.
147. Arnold L, Willoughby TL, Calkins EV, Gammon L, Eberhart G. Use of peer evaluation in the assessment of medical students. *J Med Educ.* 1981;56:35-42.
148. Marcus ER. Empathy, humanism, and the professionalization process of medical education. *Acad Med.* 1999;74:1211-5.
149. Simmons JMP, Robie PW, Kendrick SB, Schumacher S, Roberge LP. Residents' use of humanistic skills and content of resident discussions in a support group. *Am J Med Sci.* 1992;303:227-32.
150. Testerman JK, Morton KR, Loo LK, Worthley JS, Lamberton HH. The natural history of cynicism in physicians. *Acad Med.* 1996;71(10 suppl):S43-S45.
151. Satterwhite WM, Satterwhite MA, Enarson CE. Medical students' perceptions of unethical conduct at one medical school. *Acad Med.* 1998;73:529-31.
152. Eron LD. The effect of medical education on attitudes. A follow-up study. *J Med Educ.* 1958;33:25-33.
153. Wolf TM, Balson PM, Faucett JM, Randall HM. A retrospective study of attitude change during medical education. *Med Educ.* 1989;23:19-23.
154. Swick HM, Szenas P, Danoff D, Whitcomb ME. Teaching professionalism in undergraduate medical education. *JAMA.* 1999;282:830-2.
155. Brownell AKW, Côté L. Senior residents' views on the meaning of professionalism and how they learn about it. *Acad Med.* 2001;76:734-7.
156. Rowley BD, Baldwin DC Jr, Bay RC, Karpman RC. Professionalism and professional values in orthopaedics. *Clin Orthop Rel Res.* 2000;378:90-6.
157. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9 suppl):S63-S67.

### RECOMMENDATIONS TO IMPROVE FUTURE ASSESSMENT OF PROFESSIONALISM

#### Professional Behaviors as Value Clashes

126. Cohen JJ. Measuring professionalism: listening to our students. *Acad Med.* 1999;74:1010.
127. Ginsburg S, Regehr G, Hatala R, et al. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med.* 2000;75(10 suppl):S6-S11.

158. Ginsburg S, Regehr G, Stern D, Lingard L. Professionalism in context: what challenges do medical students perceive? Oral abstract, poster, and research summary presentations, 40th Annual Conference on Research in Medical Education, Group on Educational Affairs of the Association of American Medical Colleges, Washington, DC, 2001.
159. Roberts LW and the Subcommittee on Professional Attitudes and Values, Student Progress Assessment. Sequential assessment of medical student competence with respect to professional attitudes, values, and ethics. *Acad Med.* 1997;72:428-9.
160. Larkin GL. Evaluating professionalism in emergency medicine: clinical ethical competence. *Acad Emerg Med.* 1999;6:302-11.

### The Environment in Which Assessment Occurs

#### *The Institutional Stance*

161. Hafferty FW, Franks R. The hidden curriculum, ethics teaching, and the structure of

- medical education. *Acad Med.* 1994;69:861-71.
162. Hundert EM. Characteristics of the informal curriculum and trainees' ethical choices. *Acad Med.* 1996;71:624-33.
163. Markakis KM, Beckman HB, Suchman AL, Frankel RM. The path to professionalism: cultivating humanistic values and attitudes in residency training. *Acad Med.* 2000;75:141-50.
164. Stern DT. Values on call: a method for assessing the teaching of professionalism. *Acad Med.* 1996;71(10 suppl):S37-S39.
165. Stern DT. In search of the informal curriculum: when and where professional values are taught. *Acad Med.* 1998;73(10 suppl):S28-S30.
166. Stern DT. Practicing what we preach? An analysis of the curriculum of values in medical education. *Am J Med.* 1998;104:569-75.
167. Burack JH, Irby DM, Carline JD, Root RK, Larson EB. Teaching compassion and respect. *J Gen Intern Med.* 1999;14:49-55.
168. Gordon MJ. Cutting the Gordian knot: a two-part approach to the evaluation and

- professional development of residents. *Acad Med.* 1997;72:876-80.
169. Hemmer PA, Hawkins R, Jackson JL, Pangaro LN. Assessing how well three evaluation methods detect deficiencies in medical students' professionalism in two settings of an internal medicine clerkship. *Acad Med.* 2000;75:167-73.
170. Rhoton MF. Professionalism and clinical excellence among anesthesiology residents. *Acad Med.* 1994;69:313-5.
171. Van Rosendaal GMA, Jennett PA. Resistance to peer evaluation in an internal medicine residency. *Acad Med.* 1992;67:63.

#### *Administration of Assessment*

172. Hemmer PA, Hawkins R, Jackson JL, Pangaro LN. Assessing how well three evaluation methods detect deficiencies in medical students' professionalism in two settings of an internal medicine clerkship. *Acad Med.* 2000;75:167-73.
173. Gordon MJ. Cutting the Gordian knot: a two-part approach to the evaluation and professional development of residents. *Acad Med.* 1997;72:876-80.