Jack L Maatsch Visiting Scholar in Medical Education Fund

OMERAD
Office of Medical Education Research and Development
College of Human Medicine—Michigan State University

This document contains the text and references for Dr. Louise Arnold's Maatsch Scholar Award presentation at the Office of Medical Education Research and Development in the College of Human Medicine at Michigan State University in May, 2001.

A later version of this work has been accepted for publication in Academic Medicine.

Assessing Professional Behavior: Yesterday, Today, and Tomorrow©

Louise Arnold, Ph.D.
Jack L. Maatsch Lecture
May, 2001
Michigan State University College of Human Medicine

I am honored to be with you on this occasion that celebrates the work of Professor Jack L. Maatsch whose vision and expertise moved the field of assessment light years forward! He indeed was a medical educator and assessment specialist for, and ahead of, his time.

Today, I will offer you my interpretation of the research literature on assessing professional behavior. As I examined articles and books, collected over the years and identified through Med-line and Eric searches, I wore two hats. With my research hat on and considering assessment in its broadest sense (Stufflebeam *et al.*), I asked: what does this paper or book teach us about collecting information to make decisions regarding the professional behavior of learners or practitioners, educational processes or programs? With my administrator's hat on, I asked: how does this study bear on current initiatives in medical education circles to find sound methods for evaluating professional behavior in medical school, residency, and on into practice?

In my presentation I will review the concept of professionalism historically, describe major tools available for assessing professional behavior, and reflect on recommendations to improve evaluation of professionalism in the future. I trust another Jack Maatsch will emerge to push this endeavor ahead. But first, I will summarize my talk so you know where my thoughts will take you.

Summary

Thirty years ago medical educators were silent about professionalism *per se*. Today they have a delineated concept of professionalism to use as a springboard to next steps in assessing professional behavior. The current array of assessment tools is rich. But their reliability or dependability, their internal validity or credibility, and their external validity or transferability are moderate, at best. How can we strengthen our toolbox? Besides improving the psychometric properties of our measurements, we need to promote the evaluation of separate elements of professionalism. We should encourage rigorous qualitative approaches to assessment such as the analysis of critical incidents to capture professional behavior as it occurs day to day. We should combine these qualitative assessments with more quantitative measures of professional behavior, such as OSCEs that can capture competence. We need to test the hypothesis 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. Crafting new assessment strategies to assure *both* quality of and growth in

professionalism among physicians throughout the medical education continuum will be our biggest but most rewarding challenge.

The Concept of Professionalism

Yesterday. Thirty years ago, my home discipline of sociology had a crisp concept of profession. In contrast to other occupations, a profession was deemed a vocation with a body of knowledge and skills put into service for the good of others (Parsons). 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 excellance*.

In the 1970's when I entered medical education, the concepts of profession and professionalism *per se* were absent. 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 (Calkins *et al.*; Murden *et al.*; Miller *et al.*; Keck J, Arnold L *et al.*).

In the mid-1980's a major change occurred. The ABIM began its humanism project. It saw humanism as an entity consisting of respect, compassion, and integrity. It supported a number of studies for evaluating the humanism of resident physicians (ABIM, 1985). In turn, the humanism initiative led to Project Professionalism in the mid-1990's (ABIM, 1994).

Today. The concept of professionalism is clearly circumscribed with specific elements. Definitions, empirically and prospectively derived, abound. About 50% of medical schools have written criteria and specific assessment methods to assess professional behavior (Miller; Swick *et al.*) This chart summarizes the number of elements and behavioral indicators of professional behavior that schools use to evaluate their students. Professional organizations agree on the elements of professionalism. These elements include: altruism; respect for others and other humanistic qualities; honor, integrity, ethical and moral standards; accountability; excellence; and duty/advocacy[1] (ABIM, 1994; Adams *et al*. SAEM, 1998; ACGME, 1999).

Leaders in medical education concur, to a point. They add autonomy and dealing with uncertainty to the mix (Creuss *et al.*; Swick).

Authors and organizations also vary the emphasis given to some of the elements. Altruism is the lynchpin for the ABIM (1994). Duty, advocacy, service, social responsiveness are critical to the perspective of the Creusses and others (Irvine; Kimball;

[1] The ABIM attaches the following meanings to each of these elements. Altruism demands that the best interests of patients, not self-interest, guide the physician role. 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 relationships, the profession, and society. Excellence entails a commitment to exceed ordinary expectations and a commitment to life-long learning. Duty is the free acceptance of commitment to service.

Rothman). Some writers argue that autonomy and self-regulation are *passé*. Others contend these elements are more critical than ever if medicine is to remain a profession (Creuss *et al*.).

There are nuanced differences as well. Accountability contains six domains according to some (Emanuel & Emanuel). Humanism should be treated as an entity with empathy as the central concept (Gold Foundation Conference, 2001). Overlaps between elements exist (ABIM, 1985; ABIM, 1994). MSOP does not contain the concept of professionalism but does speak to its elements (MSOP 1999).

Additionally, challenges to the elements of professionalism have been recognized. Conflicts of interest, abuse of power, lack of conscientiousness – these and other challenges -- are important for assessment of professionalism.

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, the messiness, need clarification.

Measurement of Professional Behavior

Regrettably, 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 work evaluates only professional behavior, as a comprehensive entity in and of itself. Still other studies evaluate 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 will describe tools from each of these types of studies, report psychometric properties and substantive findings, and draw implications for next steps.

Research on evaluating professional behavior as part of assessing 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' assessment of their peers may be the best measure of professional and nonprofessional behavior. Peers are in frequent, close contact with each other when no one who counts is looking. Most often peer assessment relies on rating scales, but a notable study describes a successful annual nomination of top peers in a medical school graduating class (Small *et al*.). Internal consistency of peer assessment tools can be high (Arnold L. *et al*.). Inter-rater reliability is moderate (Panszi *et al*.) Yet peer assessments can suffer from a halo effect (Arnold L. *et al*.) since learners may not differentiate between peers' technical knowledge and skills and peers' professional behaviors. The relationship between peer assessments and faculty measures is weak to moderate (Arnold L. *et al*.; Panszi *et al*.) although the meaning of this finding is not clear in the absence of a gold standard of professional behavior. All too often, peers may be reluctant to assess each other (Helfer; Thomas *et al*.; Van Rosendahl & Jennett). On the other hand, peers

do offer solid information about each others' interpersonal skills (Linn *et al*.; Helfer; Schumacher). They contribute unique insight into the professional behavior of each other (Small *et al*.; Kubany).

Implications for tomorrow. Peer assessment of professional behavior 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, the elements of professional behavior they feel willing and competent to assess, and the indicators of professional behavior they believe reflect their experience (Arnold and Stern).

Studies of physicians' evaluating professional behavior as a part of clinical performance have also relied largely on rating scales. The excellent study of Ramsey *et al*. (1993) used a form containing 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, partly due to the small numbers of raters frequently used. Ratings from at least 11 physician associates of each physician subject would be needed to achieve an acceptable level of reliability, according to Ramsey *et al*. (1993). Inter-rater agreement on humanistic items is particularly low (Johnson *et al*.). However, in Ramsey's study, agreement between physician associates and nurses reached .5 corrected for attenuation. High inter-correlations across categories of behaviors abound (Saunders and Paiva; Hull; Durand; Davis). 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 (Ramsey *et al*. 1993, 1996; Arnold L. & McNeley K; Maxim & Dielman; Gough; Geertsma & Chapman). These findings suggest, then, that expert evaluators may cognitively bifurcate their perceptions of learners and peers into just two categories, without making distinct judgments among the separate elements of professional behavior although occasionally three factors have been derived from clinical performance data (Arnold L *et al*.; Benyamini *et al*.).

Implications for tomorrow. This prospect raises some unsettling issues for future work. If expert evaluators organize their perceptions into a technical knowledge/skill category and a professional behavior category, do we need, and can we obtain, ratings of the various elements of professionalism in order to certify professional behaviors across the continuum of medical education ? Perhaps not. But surely we need ratings of separate elements to guide growth in professionalism along the continuum of medical education. For that reason I would not choose a tool measuring clinical performance to assess professional behavior.

Studies exclusively focused on measuring professional behavior, by using a comprehensive definition of professionalism. These studies fall into two major categories. One type assesses *groups* of learners through surveys. The other evaluates *individuals* through critical incident techniques.

An outstanding study of the professional behavior of groups of learners tackles forthrightly the issue of whether professionalism can be measured (Arnold E. *et al*.). 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 ABIM's six elements of professionalism.

The internal reliability of the instrument was acceptable overall (.71). Factor analysis of the data yielded 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 did distinguish levels of excellence among residents in the five participating institutions. The remaining two, less reliable, factors had too few items, overlapping items, and an important altruism item that loaded on the excellence factor.

Implications for tomorrow. This study suggests that respondents can distinguish among the elements of professionalism if the tool examines only professional behaviors. Although the study just reviewed produced group scores assigned by learners who may not yet be expert raters, its encouraging results suggest this line of inquiry should be continued. Further work should try to increase reliability of the instrument across raters, ratees, and time. It should investigate whether the items reflect learners' ideas about the elements of professional behavior as well as their everyday experience.

The second line of inquiry in studies assessing professional behavior with a comprehensive definition entails the use of critical incidents. In a distinguished series of studies, Rhoton (1991;1994) 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 behavior although, it is important to note, 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 instances of unprofessional behavior in their records achieved excellent clinical performance. But those with unprofessional behavior performed poorly. Additional studies -- including those using a comprehensive definition of professionalism -- (Herman *et al*.; Rowley *et al*.; Arnold L. *et al*.; Price *et al*.; Sheehan *et al*.; Ramsey *et al*.), have found similar relationships between professional behavior and overall performance.

Other work using critical incidents to assess professional behavior entail longitudinal assessment (Loeser & Papadakis; Phelan *et al*.; Papadakis *et al*.). These longitudinal assessment programs allow 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 of class, goes to a dean who meets with the student and decides on appropriate action. The programs allow tracking of students throughout their medical school stay with the goal of remediation if necessary. 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 (Phelan *et al*.; Papdakis *et al*.).

The evaluation process itself provides for reliability since at least two reports must reach a dean before action is taken and a dean meets with the identified students for further exploration of the incident. Validity data come from case disposition. In one school, the dean found cause to take action in nine out 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 behavior. Some behavior is not quantifiable along a scale. Can you have a little bit of honesty or integrity? Quantification is difficult too because unprofessional behavior does not happen frequently. One of these programs 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 a focus only on unprofessional behaviors. Accordingly all students do not receive feedback. The absence of a report about a student's behavior is not a testament to that student's professional conduct. Faculty may be wary of the longitudinal assessment program, but they do participate.

Implications for tomorrow: On balance, these studies using critical incidents hold promise. They point up the important role of the dispassionate, disciplined reviewer of behavior -- be s/he researcher or dean. They invite qualitative analysis of unprofessional behavior through time. Usefulness of critical incident techniques may be expanded if reports about professional behavior were also sought.

Studies evaluating specific elements of professionalism: Humanism. Humanism has been evaluated through self reports, OSCEs, and rating scales. Several questionnaires eliciting self reports have been developed to characterize humanistic trends among groups of learners (Abbott; Wolff *et al*.) The questionnaire of Abbott is noteworthy. Based upon Pellegrino's concept of humanism, its psychometric properties have been thoroughly established, to good effect.

Recently an OSCE was used to see if the humanism of family medicine clerks can be predicted (Rogers & Coutts). Standardized patients used an eight-item checklist derived from a recognized scale (Hauck *et al*.) 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 value 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 "greets you warmly".

However, much of what we know about measuring professionalism stems from studies of humanism that the ABIM sparked in the 1980's. 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 to nurse and patient ratings of residents' humanism. These ratings are unreliable unless large numbers of raters are used. For example, if faculty used one global item, 20-50 observations would be required to achieve an acceptable level of reliability. If they used several items, the number of observations required would total 11. To obtain reliable ratings from nurses, between five and 20 observations would be necessary; to obtain reliable ratings from patients 20-50 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 (Ramsey *et al*.). In short, humanism of a resident depends on whom you ask.

A number of factors explain the discrepancies between ratings, in addition to the low number of raters employed in many of these studies. These factors include differential opportunities for observation. For example, in outpatient settings the difference between patient and faculty ratings of residents' humanism shrank (McLeod *et al*.). Furthermore, different raters used different criteria; that is, 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. Moreover, humanism scores given to residents varied by the gender of raters and the gender of ratees (Woolliscroft *et al*.). 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 (Klessig *et al*.). Finally, humanism scores also depended upon the ethnicity of raters (Merrill *et al*.) and the age and health status of patients (Woolliscroft *et al*.). 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 and prompt the recommendation that a profile on humanism containing information from multiple sources may be necessary and useful.

Studies evaluating specific elements of professionalism: Self assessment and the ability to self regulate, self reflect. Self assessment of professional behavior may be suspect (Ginsburg *et al*.). Afterall, self assessment of technical knowledge and skills is often inaccurate (Gordon 1991; 1992). Residents' self ratings of humanism are weakly related to others' ratings of their humanism, if at all (McLeod *et a*l., Klessig *et al*.). A new relative ranking technique appeared promising in the self assessment of interviewing skills (Regehr *et al*.). But when residents using the relative ranking technique self assessed a broad range of their clinical performances, they said they needed the least work in collegiality and team relationships; while they readily admitted they needed the

most work with their knowledge or skills (Harrington *et al*.). Further, learners are reluctant to rate themselves (Gordon 1991, 1992). 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 (Gordon 1991; 1992); 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 behavior may be incredibly difficult, work on measuring this skill with regard to professionalism needs to continue. Self assessment is a critical component of professionalism. As with measuring other elements of professionalism, identification of the conditions that could support accurate self assessment is vital.

Studies measuring other specific elements of professionalism. Today I can only touch upon the tools available to assess such elements of professionalism as altruism, duty, empathy, and ethical decision-making. Standard instruments such as personality and value inventories (Arnold L. *et al*; Davis MH; Epstein *et al*.; Magee & Hojat; Hojat *et al*.) or tests of moral reasoning (Rest) with excellent psychometric properties are available. But at least some of them might not be relevant to medical education. Our own work with empathy training (Feighny *et al*.) and studies on ethical dilemmas (Rezler *et al*.) point in that direction.

On the other hand, OSCEs that test learners' ethical reasoning, ethical behavior, and communication skills might have greater clinical relevance. Studies have found that students' performances in communication increased through time (Klamen & Williams). 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 (Pieters *et al*.). Communication OSCEs can test the ability to convey humanism to patients. Yet, OSCEs have been criticized for artificiality (Arnold, R. *et al*.). Further, any single station has low reliability (Singer *et al*.; Donnelly *et al*.). Scores are confounded by the content of the stations (Donnelly *et al*.; Hodges *et al*.), and ethical decision-making can be inextricably entwined with communication skills.

Implications for tomorrow. Standard psychological tests with outstanding psychometrics may be an excellent resource for measuring altruism, duty, empathy, and ethical and moral reasoning. 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. These stations mimic clinical situations. Since their reliability depends on the number of rating opportunities, the effort needed to generate solid tests of these elements of professional behavior by using OSCEs will be considerable.

Summary of implications regarding measures of professional behaviors. An array of tools exist for assessing professional behaviors. The reliability and validity of tools, especially rating scales, require attention if we adhere to standards of acceptable

psychometric properties of measurement. We must expend the effort. Finding or developing tools to tap the separate elements of professionalism should be a top priority. Qualitative approaches, along with OSCEs, might be helpful. The elements of an environment supportive of assessing professionalism must be ascertained. A fast easy solution eludes us.

Recommendations to Improve Assessing Professionalism in the Future

Professional behaviors express clashes between values. More than one author enjoins us to stress behaviors in assessing professionalism (Cohen; Ginsburg *et al*.). In searching for ways to improve assessment of professionalism, an innovative review by Ginsburg *et al*. notes that traditional evaluation methods rely on abstract idealized definitions that characterize people, rather than their behaviors, as unprofessional or professional. Further, these idealized traits imply that professionalism represents a set of stable traits. Several studies suggest the opposite (Sawyer; Carlo *et al*.). For example, MMPI testing of psychiatry residents identified serious personality disorders in two individuals who eventually lost their licenses for professional misconduct. Other participants also showed the same personality traits; yet no reports were lodged against them in 15 years of follow-up (Garfinkle *et al*.) Our own unpublished work using the MMPI showed a similar pattern among medical students.

Ginsburg and colleagues contend that measures of stable traits also miss the mark because they do not view professional behaviors as expressing clashes between two or more equally worthy values. Evidence certainly supports the contention that value conflicts underlie unprofessional behavior. In discussing ethical dilemmas with peers, medical students struggled with several conflicts between worthy values that led to questionable behavior (Christakis and Feudtner; Swenson and Rothstein). These included conflicts between learning medicine 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 gaining medical knowledge to become a competent physician. Faculty observed that among students, residents, and their own colleagues the values of conscientiousness and excellence could easily conflict with altruism (GEA discussion group). A survey revealed the 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 (Castellani and Wear). The survey also described some resultant, less than model, behaviors (Castellani and Wear).

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 (Hodges *et al*.). Their success led to the suggestion that OSCEs could place students in situations involving difficult value conflicts where their responses might reveal professional lapses (Hodges *et al*.).

Ginsburg and colleagues also maintain that how learners resolve the conflict between values is every bit as important as the behavior itself. Their suggestion is reminiscent of several attempts to evaluate professional behavior. These include a written follow-up to an ethics OSCE station where students explained their choice of actions (Smith *et al*.) and

a professional decisions inventory (Rezler *et al.*) in which students indicate how they would respond in a clinical scenario and then choose values to justify their response. Learners' think aloud exercises, narrative reports (Branch), responses to cases (Branch *et al*.), reflective pieces (Lingard and Haber), and focus group transcripts can be subjected to qualitative analysis to lay bear resolution of value conflicts.

Examining this process is a critical step. Not only can such a process help us to find out how a learner deals with the conflict. But also it can enable us to discover whether the learner perceives a conflict in the first place and to understand how and why the values they use might deviate from the elements of professional behavior. Here I am reminded of a report about 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 (Osborne). The distance between generations and the diversity of our students in this post-modernist world underscore the need for exploring resolution. Such exploration could reveal that some unprofessional behaviors might not reflect value clashes, but rather other etiologies.

Implications for tomorrow: Accordingly, we should test the hypothesis that measurement tools focused on professional behaviors as expressions of value conflicts will produce more reliable and valid instruments. Such tools might be useful for evaluating "routine" occasional lapses into unprofessional behavior. Research on the process of resolving value conflicts should 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 (Nash.)

Professional behaviors are context-dependent. Ginsburg and colleagues also argue that assessment in this area must also recognize the specificity of professional behaviors. Much evidence for their proposition -- that professional behaviors are depend on context – can be found in studies on ethical dilemmas (Rezler *et al*.) where values that the same individuals brought to bear in taking actions varied across the scenarios presented. Our own work illustrates the situational specificity of unprofessional behavior since the frequency of peers reporting negative behaviors of their colleagues was a function of the quality of leadership on the health care team (Arnold L. *et al*.). That is, groups with leaders who were physically absent or who used laissez faire techniques had a greater frequency of negative peer reports than 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' ability to interact with patients empathically and altruistically (Marcus). 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 (Simmons *et al*.) Cross-sectional, longitudinal, and retrospective studies of cynicism and humanism illustrate similar ups

and downs. Students felt they grew more cynical during medical school but also more interested in and helpful with patients (Wolff *et al*.). Medical students are most cynical, while residents and especially faculty are less so (Testerman *et al*.). Professional behaviors appropriate for small groups in the early years of medical school have been delineated (Bienenstock).

From these results flows the suggestion that what physicians need to learn and thus what needs to be assessed regarding professional behavior will vary according to career stage. Swick *et al*. 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. By relating the elements of professionalism to medical students' main tasks of gaining and applying knowledge and skills to deal with patient problems, another investigator selected much the same set of elements but added self regulation and accountability for self and peers (Brownell). On the other hand, educators who subscribe to the effectiveness of anticipatory socialization should argue that all of the elements of professional behavior should apply to medical students. For residents, the ACGME has specified the elements of professionalism presented at the beginning of this talk. Yet residents themselves have defined professionalism as entailing only competence (Brownell and Cote). 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 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. In contrast, approaches to longitudinal assessment of the ethical development of medical students (Roberts *et al*.) and residents (Larkin) use the know, can, do model.

Implications for tomorrow. Two issues await resolution. Should all or only some elements of professionalism be assessed at different stages of a medical career? Should all levels or only some levels of assessment be used during the various stages of a medical career? 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 even 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, we can not neglect the know and can levels of the pyramid.

Environment in which assessment occurs. The researcher in me says "ok, we have our concept, we need to improve our tools, and we need to find out if the new approaches to assessing professional behavior really do help us. We are ready to go to work! But my experience in the medical school prompts me to say that won't be enough. All our elegant theorizing and precise measurement will not get us to where we need to be unless

we 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.

The institutional stance. Theoretical and empirical studies rivet our attention on the hidden and informal curriculum (Hafferty; Hundert). A great deal of teaching about professional values occurs outside of scheduled class-time, in the informal curriculum, when faculty are absent (Stern). Further, only some of the "teaching" there is congruent with the announced professional values of an institution. According to one study, the taught curriculum emphasized industriousness of learners, while the professed curriculum was silent on that matter (Stern). The taught curriculum also spoke to the burden of service and interprofessional disrespect rather than their opposites. Only if we know the lessons of the informal curriculum 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 (Burack; Bienenstock; Gordon; Hemmer *et al.*; Rhoton; Van Rosendahl and Jennett). To do that, we need to know the stance of the informal culture on assessing professional behavior: just how important is it, 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. We need to know whether there is courage to follow through with discipline if necessary and whether the highest levels of administration support or pay lip service to assessment of professional behavior.

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 transgression, sickness, and deviance? Is it done in the spirit of justice and the social good or only for the good of the individual? Clarity concerning the purpose of the assessment must be achieved and communicated. How much of the assessment is formative for guidance, growth, and striving toward the ideal; how much is done for summative reasons? What consequences does the assessment hold for counseling, grades, and promotion? Is the environment safe for assessment of professional development? What constitutes safety for learners and faculty? Is the assessment anonymous, confidential, or signed? 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 comments section of an evaluation form (Hemmer *et al.*). Who does the evaluation; do the most vulnerable people in the system have input? Who receives the evaluation; a credible fair reviewer? Finally, do all participants receive education in the assessment of professional behavior?

Noting the dissatisfaction with evaluation systems in residencies, Gordon 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, peer and faculty coaching. Faculty would insist only that residents would participate in good faith. Results of our work in progress on self assessment suggest that this insistence would be necessary.

Implications for tomorrow. The acceptability and efficacy of Gordon's proposal should be studied in the context of professional behavior. How it could be adapted to undergraduate medical education should be examined.

Summary

Thirty years ago medical educators were silent about professionalism *per se*. Today they have a delineated concept of professionalism to use as a springboard to next steps in assessing professional behavior. The current array of assessment tools is rich, but we need to strengthen them to achieve dependable, credible, and transferable measurements. We need to promote the evaluation of separate elements of professionalism. We should encourage rigorous qualitative approaches to assessment such as the analysis of critical incidents to capture professional behavior as it occurs day to day. We should combine the qualitatively derived insights with more quantitative measures of professional behavior such as OSCEs that can capture competence. We need to test the hypothesis 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. Crafting new assessment strategies to assure *both* quality of and growth in professionalism among physicians throughout the medical education continuum will be our biggest but most rewarding challenge.

Bibliography

Introduction

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

The Concept of Professionalism

ABIM. *A Guide to Awareness and Evaluation of Humanistic Qualities in the Internist*. Portland, OR: American Board of Internal Medicine, 1985

ABIM. *Project Professionalism*. Philadelphia, PA; American Board of Internal Medicine, 1994.

Adams J *et al*. Professionalism in Emergency Medicine. *Acad. Emergency Med*. 1998*:*5:1193-1199.

Accreditation Council for Graduate Medical Education. ACGME Outcome Project. Enhancing Residency Education through Outcomes Assessment; General Competencies. Version 1.3, 1999.

Calkins EV *et al*. Impact on Admission to a School of Medicine with An Innovation in Selection Procedures. *Psychol Reports*. 1974;35:1135-1142.

Creuss RL, Creuss SR. Teaching Medicine as a Profession in the Service of Healing. *Acad. Med*. 1997;72:941-952.

Creuss RL, Cruess SR, Johnston SE. Renewing Professionalism: An Opportunity for Medicine. *Acad. Med*. 1999;74:878-884.

Creuss RL-196., Cruess SR, Johnston SE. Professionalism and Medicine's Social Contract. *J Bone and Joint Surgery*. 2000;82A:1189-1194.

Creuss RL, Cruess SR, Johnson SE. Professionalism: An Ideal to Be Sustained. *Lancet.* 2000;356:156-159.

Emanuel EJ, Emanuel LL. What Is Accountability in Health Care? *Ann. Intern. Med.* 1996;124:299-239.

Gibson DD *et al*. Creating a Culture of Professionalism. *Acad. Med*. 2000:75;509.

Gold Foundation. Overcoming the Barriers to Sustaining Humanism in Medicine: Influencing the Culture through a Humanism Honor Society. Invitational Conference. New Jersey, March, 2001.

Hunt DD. Functional and Dysfunctional Characteristics of the Prevailing Model of Clinical Evaluation systems in North American Medical Schools. *Acad. Med.* 1992;67:254-259.

Hunt DD. Types of Problem Students Encountered by Clinical Teachers on Clerkships. *Med. Educ*. 1989;23:14-18.

Ingelfinger FJ. Arrogance. *N Engl J Med.* 1980;303:1507-1511.

Irvine D. The Performance of Doctors: The New Professionalism. *Lancet.* 1999;353:1174-1177.

Kassirer JP. Medicine at Center Stage. *N Engl J Med*. 1992;328:1268-1269.

Keck J, Arnold L. et al. Efficacy of Cognitive/Noncognitive Measures in Predicting Resident Physician Performance. *J Med. Educ.* 1979;54:759-65.

Kimball HR. Medical Professionalism: At A Crossroad. 11th Annual Coggeshall Memorial Lecture. University of Chicago, 2000.

Miller GD *et al*. Noncognitive Criteria for Assessing Students in North American Medical Schools. *Acad. Med*. 1989;64:42-44.

Murden R. et al. Academic and Personal Predictors of Clinical Success in Medical School. *J Med. Educ*. 1978;53:711-719.
Parsons T. *The Social System*. Glencoe: Ill: The Free Press, 1951.
Rothman DJ. Medical Professionalism –Focusing on the Real Issues. *N Engl J Med.* 2000;342:1284-1286.
Southwick F. Who Was Caring for Mary? *Ann. Int. Med.* 1993;118:146-148.
Swick HM. Academic Medicine Must Deal with the Clash of Business and Professional Values. *Acad. Med*. 1998;73:751-755.
Swick HM *et al*. Teaching Professionalism in Undergraduate Medical Education. *JAMA*. 1999;282:830-832.
Swick HM. Toward a Normative Definition of Medical Professionalism. *Acad. Med*. 2000;75:612-616.
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*. 1999:74:13-18.

Studies of Clinical Performance

Learners' Peer Assessment

Arnold L *et al*. Use of Peer Evaluation in the Assessment of Medical Students. *J Med Educ*. 1981;65:35-41.
Arnold L, Stern DT. Towards Assessing Professional Behaviors of Medical Students through Peer Evaluation. Working Papers, 2001.
Helfer RE. Peer Evaluation: Its Potential Usefulness in Medical Education. *Br J Med Educ*. 1972;6:2245-231.
Linn BS, Arostegui M, Zeppa R. Performance Rating Scale for Peer and Self Assessment. *Br J Med Educ*. 1975;9:98-101.
Kubany AJ. Use of Sociometric Peer Nominations in Medical Education Research. *J Appl Psychol*. 1957;41:389-394.
Panszi S *et al*. What Do Peers Know about Professionalism? Research in Medical Education Conference, Group on Educational Affairs, AAMC Annual Meeting, Chicago, IL, 2000.
Schumacher CFJ. A Factor Analytic Study of Various Criteria of Medical Examinations. *J Med Educ*. 1964; 39:192
Small PA Jr. *et al*. Issues in Medical Education: Basic Problems and Potential Solutions. *Acad. Med*. 1993;68: S89-S93.
Thomas PA *et al*. A Pilot Study of Peer Review in Residency Training. *J. Gen Intern Med*. 1999;14:551-554.
Van Rosendahl GMA, Jennett PA. Resistance to Peer Evaluation in an Internal Medicine Residency. *Acad. Med*. 1992:67:63.

Physicians Assessing Peers, Residents, Medical Students

Arnold L. *et al*. Towards Understanding the Clinical Performance of Physicians. *J Med. Educ.* 1984:59;591-594.
Benyamini K *et al*. How Do Supervising Doctors Construe the Medical Student in Clinical Training? *Med. Educ*. 1987;21:410-418.

Dawson Saunders B, Paiva R. The Validity of Clerkship Performance Evaluations. *Med Educ*. 1986;20:240-245.
Davis JK *et al*. Inter-rater Agreement and Predictive Validity of Faculty Ratings of Pediatric Residents. *J Med. Educ*. 1986;61:901-905.
Durand RP *et al*. Teachers' Perceptions Concerning the Relative Values of Personal and Clinical Characteristics and Their Influence on the Assignment of Students' Clinical Grades. *Med. Ed*. 1988;22:335-341..
Gertsma RJ, Chapman JE. The evaluation of medical students. *J Med Educ*. 1967;42:938-948.
Gough HG *et al*. Evaluation of Performance in Medical Training. *J Med Educ*. 1964;39:679-692.
Hull AL *et al*. Validity of Three Clinical Performance Assessment of Internal Medicine Clerks. *Acad. Med*. 1995;70:517-522.
Johnson D, Cujec B. Comparison of Self, Nurse, and Physician Assessment of Residents Rotating through an Intensive Care Unit. *Crit Care Med*. 1998:26:1811-1816.
Littlefield J *et al*. Literature Assessment: The Validity of Assessing Physician Professional Behavior. Presentation at the Fall Meeting of the Society of Directors of Research in Medical Education, Washington, D.C., November 1996.
Maxim BR, Dielman TE. Dimensionality, Internal Consistency, and Inter-rater Reliability of Clinical Performance Ratings. *Med. Educ*. 1987;21:130-137.
Ramsey PG *et al*. Use of Peer Ratings to Evaluate Physician Performance. *JAMA* . 1993:1655-1660.
Ramsey PG *et al*. Feasibility of Hospital-based Use of Peer Ratings to Evaluate the Performances of Practicing Physicians. *Acad. Med*. 1996;71:364-370.

Studies of Professionalism Using A Comprehensive Definition

Arnold EL *et al*. Can Profesionalism Be Measured? The Development of a Scale for Use in the Medical Education Environment. *Acad. Med*. 1998;73:1119-1121.
Herman M *et al*. Validity and Importance of Low Ratings Given Medical Graduates in Noncognitive Areas. *J. Med. Educ*. 58:8387-843, 1983.
Lavin B, Pangaro L. Internship Ratings as a Validity Outcome Measure for an Evaluation system to Identify Inadequate Clerkship Performance. *Acad. Med*. 1998;73:998-1002.
Littlefield JH. A Statement of Values. *Pedagogue.* 1995;6:2.
Loeser H, Papadakis M. Promoting and Assessing Professionalism in the First Two Years of Medical School. *Acad Med*. 2000;75:509-510.
Phelan S *et al*. Evaluation of the Non-cognitive Professional Traits of Medical Students. *Acad. Med*. 1993:68:799-803.
Papadakis MA et al. A Strategy for the Detection and Evaluation of Unprofessional Behavior in Medical Students. *Acad. Med*. 1999;74:980-990.
Price PB *et al*. Measurement of Physician Performance. *J Med Educ*. 1964;39:203-211.
Rhoton MF. A New Method to Evaluate Clinical Performance and Critical Incidents in Anesthesia: Quantification of Daily Comments by Teachers. *Med. Educ*. 1989;23: 280-289.
Rhoton MF *et al*. Influence of Anesthesiology Residents' Noncognitive Skills on the Occurrence of Critical Incidents and the Residents' Overall Clinical Performances. *Acad. Med*. 1991;66:359-361.

Rhoton MF. Professionalism and Clinical Excellence among Anesthesiology Residents. *Acad. Med*. 1994;69:313-315.
Rowley BD *et al*. Can Professional Values Be Taught? A Look at Residency Training. *Clin Orthopaedics and Related Research*. 2000:378:110-114.
Sheehan TJ *et al*. Moral Judgement as a Predictor of Clinical Performance. *Eval. Health Prof*. 1985;8:379-400.

Studies of Separate Elements of Professionalism

Humanism

Abbott LC. A Study of Humanism in Family Physicians. *J Fam Practice.* 1982;16: 1141-1146.
Beckman H. *et al*. Measurement and Improvement of Humanistic Skills in First-year Trainees. *J Gen Intern Med*. 1990;5:422-45.
Blurton RR, Mazzaferri EL. Assessment of Interpersonal Skills and Humanistic Qualities in Medical Residents. *J. Med. Educ*. 1985;60:648-650.
Butterfield PS, Mazzaferri EL. A New Rating Form for Use by Nurses in Assessing Residents' Humanistic Behavior. *J Gen Intern Med* 1991;6:155-161.
Hauck FR *et al*. Patient Perceptions of Humanism in Physicians: Effects on Positive Health Behaviors. *Fam. Med*. 1990;22:447-452.
Klesssig J *et al*. Evaluating Humanistic Attributes of Internal Medicine Residents. *J Gen Intern Med*. 1989;4:514-521.
Linn LS *et al*. Measuring Physicians' Humanistic Attitudes, Values, and Behaviors. *Med. Care*. 1987;25:504-513.
Linn LS, Cope DW. Sociodemographic and Premedical School Factors Related To Postgraduate Physicians' Humanistic Performance. *West J Med*. 1987; 147;99-103,
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
McLeod PJ *et al*. Faculty Ratings of Resident Humanism Predict Patient Satisfaction Ratings in Ambulatory Medical Clinics. *J Gen Intern Med*. 1994;9:321-326.
Merrill JM *et al*. Measuring Humanism in Medical Residents. *South Med J*. 1986;79: 141-144.
Merrill JM *et al*. Culture as a Determinant of Humanistic Traits in Medical Residents. *South Med J*. 1987;80:233-236.
Rogers JC, Coutts L. Do Students' Attitudes during Preclinical Years Predict Their Humanism as Clerkship Students? *Acad. Med*. 2000;75:S74-77.
Weaver MJ *et al*. A Questionnaire for Patients' Evaluations of Their Physicians' Humanistic Behaviors. *J Gen Intern Med*. 1993;8:135-139.
Woolliscroft *et al*. Resident-Patient Interactions: The Humanistic Qualities of Internal Medicine Residents Assessed by Patients, Attending Physicians, Program Supervisors, and Nurses. *Acad. Med*. 1994;69:216-224.

Self Assessment

Arnold L *et al*. Self-Evaluation: A Longitudinal Perspective. *J Med. Educ* 1985;60:21-28.
Gordon MJ. A Review of the Validity and Accuracy of Self-assessment in Health

Professions Training. *Acad. Med*. 1991;66:762-769.
Gordon MJ. Self-assessment Programs and Their Implications for Health Professions Training. *Acad. Med*. 1992;67:672-279.
Harrington JP *et al*. Applying a Relative Ranking Model to the Self Assessment of Extended Performances. *Adv. Health Sci. Educ*. 1997;2:17-25.
Regehr G *et al*. Measuring Self-Assessment Skills: An Innovative Relative Ranking Model. *Acad. Med*. 1996;71:S52-S54.

Altruism, Duty, Empathy, Ethical Decision-Making, Communication Skills

AAMC. Matriculating Student Questionnaire and Medical School Graduation Questionnaire.
Arnold L *et al*. Professional and Personal Characteristics of Graduates as Outcomes of Differences between Combined Baccalaureate-M.D. Degree Programs. *Acad. Med*. 1996;71:S64-S66.
Arnold RM. Assessing Competence in Clinical Ethics: Are We Measuring the Right Behaviors? *J Gen Intern Med*. 1993;8;52-54.
Batenburg V *et al*. Are Professional Attitudes Related to Gender and Medical Specialty? *Med Educ*. 1999;22:489-492.
Carlo G. et al. The Altruistic Personality: In What Contexts Is It Apparent? *J. Pers. Soc. Psychol*. 1991;61:450-458.
Clack GB, Head JO. Gender Differences in Medical Graduates' Assessment of Their Personal Attributes. *Med. Educ*. 1999;33:101-105.
Crandall SJ *et al*. Medical Students' Attitudes toward Providing Care for the Underserved. *JAMA*. 1993;269:2519-2523.
Davis MH. A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog of Selected Documents in Psychol*. 1980;10:85.
De Monchy C. Professional Attitudes of Doctors and Medical Teaching. *Med Teacher*. 1992;14:327-331.
Donnelly MB *et al*. Assessment of Residents' Interpersonal Skills by Faculty Proctors and Standardized Patients: A Psychometric Analysis. *Acad Med*. 2000;75:S93-S95.
Dornbush RL *et al*. Medical School, Psychosocial Attitudes, and Gender. *JAMWA*. 1991;45:150-152.
Dornbush RL et al. Maintenance of Psychosocial Attitudes in Medical Students. *Soc Science and Med*. 1985;21:107-109.
Eckenfels EJ. Contemporary Medical Students' Quest for Self-fulfillment through Community Service. *Acad. Med*. 1997;72:1043-1050.
Eliason BC, Schubot DB. Personal Values of Exemplary Family Physicians: Implications for Professional Satisfaction in Family Medicine. *J Fam Prac*. 1995;41:251-256 (1995).
Epstein L *et al*. On Becoming a Physician: Perspectives of Students in Combined Baccalaureate-M.D. Degree Programs. *Teaching and Learning in Medicine: An International Journal*. 1994;6:102-107.
Evans BJ et al. Measuring Medical Students' Empathy Skills. *Br J Med Psychol*. 1993;66:121-122.
Feighny KM, Arnold L *et al*. In Pursuit of Empathy and Its Relation to Physician Communication Skills: Multidimensional Empathy Training for Medical Students

Annals Behavioral Science and Med Educ. 1998;5:13-21.
Finocchio LJ *et al*. Professional Competencies in the Changing Health Care System: Physicians' Views on the Importance and Adequacy of Formal Training in Medical School. *Acad. Med*. 1995;70:1023-1028.
Hobfall S*E et al*. Personality Factors as Predictors of Medical Students' Performance. *Med. Educ*. 1982;16:251-258.
Hodges B *et al*. Evaluating Communication Skills in the OSCE Format: Reliability and Generalizability. *Med. Edu*c. 1996;30:38-43.
Hojat M *et al*. Medical Students' Personal Values and Their Career Choices a Quarter-Century Later. *Psychol. Rep*. 1998;83:243-248.
Klamen DL, Williams RG. The Effect of Medical Education on Students' Patient-Satisfaction Ratings. *Acad. Med*. 1997;72:57-61.
Krebs D. Empathy and Altruism. *J Pers Soc Psychol*. 1975: 1134-1146.
Krol D *et al*. Factors Influencing the Career Choice of Physicians Training at Yale-New Haven Hospital from 1929 through 1994. *Acad. Med*. 1998;73:313-317.
Magee M, Hojat M. Personality Profiles of Male and Female Positive Role Models in Medicine. *Psychol Rep*. 1998;82:547-559.
Paris MJ. Attitudes of Medical Students and Health Professionals towards People with Disabilities. *Arch Phys Med Rehabil*. 1993;74:818-825.
Pieters HM *et al*. Simulated Patients in Assessing Consultation Skills of Trainees in General Practice Vocational Training: A Validity Study. *Med. Educ*. 1994;38:226-233.
Ramsey PG. Evaluation of Humanistic Qualities and Communication Skills. *J Gen Intern Med*. 1993:8:154.
Rest JR. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press, 1979.
Rest JR. *Moral Development: Advances in Research and Theory*. New York: Praeger, 1986.
Rezler AG *et al*. Assessment of Ethical Decisions and Values. *Med Educ*. 1992;26:7-16.
Schnabel GK *et al*. The Assessment of Interpersonal Skills Using Standardized Patients. *Acad Med*. 1991;66:S34-36.
Singer PA *et al*. Performance-based Assessment of Clinical Ethics Using an Objective Structured Clinical Examination. *Acad. Med*. 1996;71:495-498.
Sawyer J. The Altruism Scale: A Measure of Cooperative, Individualistic, and Competitive Interpersonal Orientation. *Am J Soc*. 1966;71:407-416l

Recommendations

Ginsburg S *et al*. Context, Conflict, and Resolution: A New Conceptual Framework for Evaluating Professionalism. *Acad. Med*. 2000;75: S6-S11.

<u>Professional Behavior as Value Clashes</u>

Castellani B, Wear D. Physician Views on Practicing Professionalism in the Corporate Age. *Qualitative Health Research*. 2000:10:490-506.
Clark PG. Values in Health Care Professional Socialization: Implications for Geriatric Education in Interdisciplinary Teamwork. *The Gerontologist*. 1997;37:441-

451.
Cohen JJ. Measuring Professionalism: Listening to Our Students. *Acad Med*. 1999; 74:1010.
Christakis DA, Feudtner C. Ethics in a Short White Coat: The Ethical Dilemmas That Medical Students Confront. *Acad. Med*. 1993; 68:249-254.
Feudtner C, Christakis DA, Christakis NA. Do Clinical Clerks Suffer Ethical Erosion? Students' Perceptions of Their Ethical Environment and Personal Development. *Acad. Med*. 1994;69:670-679.
Garfinkle *et al*. Boundary Violations and Personality Traits among Psychiatrists. *Can J. Psych*. 1997;42:758-763.
Nash RJ. *Answering the Virtuecrats: A Moral Conversation on Character Education*. New York: Teachers College Press, 1997.
Smith SR *et al*. Performance-based Assessment of Moral Reasoning and Ethical Judgment among Medical Students. *Acad. Med*. 1994;69:381-386.
Swenson SL, Rothstein JA. Navigating the Wards: Teaching Medical Students to Use Their Moral Compasses. *Acad. Med*. 1996;71:591-594.

<u>Professional Behavior as Context Dependent and Stages of a Medical Career</u>

Bienenstock D *et al*. Defining and Evaluating Professional Behavior in the MD Program: 1988-1995. *Pedagogu*e. 1995;6:1-11.
Branch WT et al. Becoming a Doctor: Critical Incident Reports from Third-Year Medical Students. *N Engl J Med*. 1993;329:1130-1132.
Branch WT *et al*. A New Educational Approach for Supporting the Professional Development of Third-Year Medical Students. *J Gen Intern Med*. 1995;10:691-694.
Branch WT. Deconstructing the White Coat. *Ann Intern Med*. 1998;129:740-741.
Brownell AKW. e-mail communication GEA Professionalism Project. January 17, 2001.
Brownell AKW, Cote L. Senior Residents' Views on the Meaning of Professionalism and How They Learn about It. *Acad Med*. 2001;76 July.
Eron L. The Effect of Medical Education on Attitudes. A Follow-up Study. *J Med Educ*. 1958;33:25-33.
Larkin GL. Evaluating Professionalism in Emergency Medicine: Clinical Ethical Competence. *Acad Emergency Med*. 1999;6:302-311.
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:S124-S127.
Marcus ER. Empathy, Humanism, and the Professionalization Process of Medical Education. *Acad Med*. 1999;74:1211-1215.
Miller GE. The Assessment of Clinical Skills/Competence/Performance. *Acad. Med*. 1990;65:S63-S7l.
O"Donnell J. e-mail communication GEA Professionalism Project. January, 2001.
Osborne E. Punishment: A Story for Medical Education. *Acad. Med*. 2000;75:241-244.
Roberts LW *et al*. Sequential Assessment of Medical Student Competence with Respect to Professional Attitudes, Values, and Ethics. *Acad. Med*. 1997;72:428-429.
Satterwhite WM *et al*. Medical Students' Perceptions of Unethical Conduct at One Medical School. *Acad Med*. 1998;73:529-531.
Simmons JMP *et al*. Residents' Use of Humanistic Skills and Content of Resident

Discussions in a Support Group. *Am J Med Sci*. 1992;303:227-232.
Swick *et al*. Teaching Professionalism in Undergraduate Medical Education. *JAMA*. 1999;282:830-832.
Testerman JK *et al*. The Natural History of Cynicism in Physicians. *Acad. Med*. 1996;71:S43-S45.
Wolf TM *et al*. A Retrospective Study of Attitude Change during Medical Education. *Med. Educ*. 1989;23:19-23.

Environment

Burack JH *et al*. Teaching Compassion and Respect. *J Gen Intern Med*. 1999;49:14-55.
Gordon MJ. Cutting the Gordian Knot: A Two-part Approach to the Evaluation and Professional Development of Residents. *Acad. Med*. 1997;72:876-880.
Hafferty FW, Franks R. The Hidden Curriculum, Ethics Teaching, and the Structure of Medial Education. *Acad. Med*. 1994;69:861-871.
Hemmer PA *et al*. Assessing How Well Three Evaluation Methods Detecting Deficiencies in Medical Students' Professionalism in Two Settings of An Internal Medicine Clerkship. *Acad Med*. 2000;75:1677-173.
Hundert EM. Characteristics of the Informal Curriculum and Trainees' Ethical Choices. *Acad. Med*. 1996;71:624-633.
Markakis KM *et al*. The Path to Professionalism: Cultivating Humanistic Values and Attitudes in Residency Training. *Acad. Med*. 2000;75:141-150.
Stern DT. Values on Call: A Method for Assessing the Teaching of Professionalism. *Acad. Med.* 1996;71:S37-39.
Stern DT. Practicing What We Preach? An Analysis of the Curriculum of Values in Medical Education. *Am J Med*. 1998;104;569-575.
Stern DT. In Search of the informal Curriculum: When and Where Professional Values Are Taught. *Acad. Med*. 1998;73:S28-S30.