The Metrics of Medical Professionalism: Trends in Education, Licensure, and Certification

Humanistic Qualities/Professionalism

Unsatisfactory			Satisfactory			Superior		
1	2	3	4	5	6	7	8	9

Internal Medicine Grand Rounds
6 July 2006

James M. Wagner, MD, MSc
Assoc. Professor, Internal Medicine
Assoc. Dean for Student Affairs

Dr. Wagner has no financial arrangements or affiliations with companies whose products or equipment referenced in this CME activity. He will not be discussing "off label" uses in this presentation.

Introduction

Americans are justifiably proud if their health-care system and physicians. The system educates its physicians well and is capable of offering high quality medical care and dizzying state-of-the-art technology to effect relief commonly and cures increasingly. The physicians are talented, intelligent, and committed to the welfare of their patients. The American health care system and its physicians are also devoted to continuous quality improvement, which requires the uncomfortable procedure of occasionally focusing on what is done poorly to grasp the problem and better align the physicians and our system.

The measurement of professionalism is a new concept. It is new in that before the 1970s cognitive knowledge was equated with professional competency. As technology advanced, physicians became more able to help their patients while at the same time became distanced from their patients. In the 1970s, articles about "humanism" in medicine became more prevalent in the literature. The American Board of Internal Medicine (ABIM) took the lead in the 1980s by defining and measuring humanism in their "Humanism Project". Eventually, "humanism" became the "professionalism" movement, and ABIM's "Humanism Project" became in the 1990's "Project Professionalism"[1].

Today's discussion departs from the scientific realm characteristic of this venue. The consideration of the concept of professionalism involves non-science topics like philosophy, ethics, and the law. Although the science of the metrics of professionalism is improving, it will never reach the level of analysis of molecular biology. In fact, it will be argued that it should not reach that level.

The most common experience measuring professionalism is by a global rating, typically of a student or resident by an attending physician. This experience can be bewildering, as there is a dizzying array of facets of professionalism, many of which are quite subjective. The challenge of combining these facets into a single scale is daunting. This can lead to a perception that the measurement of professionalism is impossible.

In addition, the contemplation of professionalism can be emotional. On one hand, most physicians choose their career because of a passionate commitment to serving mankind, one of the basic tenets of professionalism. On the other hand, most physicians can easily recall an example of their own uninspiring behavior. This can produce strong, emotional conflicts, which may lead to reluctance to consider issues of professionalism in themselves or others. This will be revisited later.

As the assessment of professionalism starts to provide data, the results are disturbing. As the toolkit we use to assess professionalism grows and the quality of the tools improves, an understanding is developing that there is room for improvement in how professionalism is taught and encouraged.

The main points of this Grand Rounds can be summarized as follows. An attempt will be made to convince the audience that professionalism should and can be measured, but we can do better.

1) We must shift our thinking about professionalism.
   a) Shift from measuring values to behaviors.
   b) Shift from a reductionist toward an expansionist perspective
   c) Shift from a summative toward a formative evaluation
2) We each are an important part of the teaching/evaluation of professionalism
   a) Our formative evaluations should be frequent and constructive
   b) Our summative evaluations should be unhindered by legal concerns
   c) We must be good role models

The first section of this presentation focuses on uncomfortable facts regarding the current state of professionalism. The second section will offer solace and hope through exploring what is being done about this poor performance and what the future holds.

Should Professionalism be measured?

Our profession provides a vital service to society. If the quality of this product has been shown to be wanting, this challenges the basic tenets of being a physician. Unfortunately, there is evidence for problems with professionalism along the entire spectrum of medical careers from training to retirement.

Is there a Problem during Medical Training?

There are several problems that indicate a problem with professionalism during training. The problems are apparent in the rates of cheating during the basic science years before contact with patients, and new problems of cynicism and retarded moral development arise during the clinical years. In addition, the press received by outrageous, criminal actions of residents and physicians are painful reminders of the problem.

Cheating

Cheating represents a violation of a fundamental principle of professionalism: honesty. In undergraduate and graduate medical education, several studies have demonstrated high rates of academic dishonesty when taking tests. The first survey of medical students, done in 1980 on all levels of medical students at two Chicago schools, revealed that 58.2% had cheated at least once in medical school[2].

Cynicism

Although a rise in cynicism has been documented and discussed for decades, Becker's landmark sociologic study of medical education, *Boys in White*[3], was the first to document in detail how medical students make the transition from laypersons to physicians. Becker contends that although they maintain their idealism, cynicism is introduced in reaction to the stressors necessitated by the process of medical education. That is, passing tests and courses, and dealing with demanding hours, difficult patients, and challenging supervisors produces an environment that nurtures cynicism.

The rise in prevalence of cynicism during medical school has been documented in several studies. Eron reported scores on a cynicism scale rose when comparing freshmen to senior medical students, while scores fell in law and nursing students[4]. Wolf reported that cynicism was the most profound change in attitude among ten positive and negative attitudes surveyed among graduating Louisiana State University medical students; in addition to being the biggest change, it was reported by 76% of respondents[5]. This prevalence has been confirmed in other studies[6].

The rise in cynicism is disturbing in itself, but its consequences are also concerning. Wear writes,
> "… although such humor may have a significant psychological benefit to those who use it, it may compromise 'the aggregate character of the medical enterprise."

"Ethical Erosion"

Beyond cheating on tests and cynicism, there are problems with professionalism in the clinical realm. Several studies have documented a rise in cynicism and "ethical erosion" during medical school[3, 7, 8].

Many have written of a "hidden curriculum" that teaches students and residents cynicism and other principles unbecoming of a medical professional[9, 10]. Feudtner reported more detail of these stressors and how they pertain to professionalism. In a 1994 survey of 665 clinical students at six Pennsylvanian medical schools, 62% felt that at least some of their ethical principles had been eroded or lost. 61% had witnessed what they believed to be unethical behavior, and 58% believed they had done something unethical themselves. Examples of unethical behavior included derogatory humor at the expense or patients (reported by 98% of students), performing unnecessary forceps deliveries "for practice", and sedating a patient who declined medications for valid reasons for the purpose of administering them intravenously.

Students who had witnessed unethical behavior were more likely to have acted improperly themselves. Students who had behaved poorly stated they had done so either for fear of a poor evaluation or to fit in with the team; these students were twice a likely to report a belief that an erosion of their ethics.[7]

Swango

Probably the most egregious and widely-known example of unprofessional conduct was published in the book *Blind Eye*[11]. A Pulitzer Prize winning author wrote about Dr. Michael Swango's medical career and consequent FBI investigation. Dr. Swango is suspected in up to 60 deaths of patients and associates and many more near deaths, making him one of the most prolific serial killers in history. In 1984, after one year of internship at Ohio State University Hospital, he was terminated after he confessed to administering a lethal dose of potassium chloride to a patient. In 1985, he was convicted of aggravated battery for poisoning co-workers with arsenic; he spent 5 years in prison.

Soon after his release, he completed several inaccurate applications and was eventually admitted to an internal medicine residency at SUNY Stony Brook.

He was convicted of fraud in 1998 for gaining admission to a residency at the State University of New York under false pretenses. He was eligible for parole in 2000, but the FBI investigation of the deaths of four patients while Swango was a resident prompted a guilty plea in 2000, resulting in four consecutive life sentences with no possibility of parole[12].

Two themes predominate the book. First, several supervisors from several campuses missed many opportunities to end Swango's career and killing spree. Second, the medical profession failed to detect and track egregious actions of physicians.

From the inside flap: "Blind Eye shows us the danger we face in a hospital system that too often puts appearances, reputation and potential liability ahead of patients' welfare – and tells us what needs to be done to stop it. It exposes the alarming failure of a national policy that is supposed to monitor incompetence and misconduct in the medical profession, as well as the weakness of the federal legislation that is intended to protect the public."

The public must question, "If the system cannot catch a serial killer, how well does it do with less egregious instances of unprofessional conduct?"

Is There a Problem in Medical Practice?

The unfortunate answer is again "yes". The problem is apparent in data reported by state medical boards and health services research. Health services research has demonstrated an unsatisfactory overall quality of care, leading to a costly medical error rate. In addition, there is good evidence to support that the quality of care delivered decreases as physicians progress through their careers. This forecasts a general decrease in overall quality as the mean age of our population (and its physicians) rises.

State Medical Board Data

Although only 0.5% of physicians are disciplined by state medical boards nationwide, it is estimated that 3-5% of physicians behave in an unprofessional manner. According to the Texas State Medical Board, they receive ten-fold more complaints related to interpersonal skills than lack of knowledge or poor technical skill worthy of an investigation[13]. This is also true nationally[14]. In a study of an innovative patient complaint system at Vanderbilt University, Hickson found that 5% of physicians were associated with one third of all patient complaints submitted[15].

Studies of Quality of Care

Several recent studies illustrate disturbing facts regarding quality of care provided in the United States. The Institute of Medicine published a report in 2000 entitled "To Err is Human[16]", which is credited with sparking the quality movement[17]. The most widely quoted statement from this report is "At least 44,000 people, and perhaps as many as 98,000 people, die in hospitals each year as a result of medical errors that could have

been prevented, according to estimates from two major studies." The estimated cost of these errors is between $17 billion and $29 billion per year. The public response to this report was profound and sustained[17]. In addition, health insurance companies are very interested in maximizing the quality of care provided.

The point of this report in the context of professionalism is that the problem is not an unprofessional doctor responsible for the majority of these mistakes; it is a system problem. In the words of the report, "People working in health care are among the most educated and dedicated workforce in any industry. The problem is not bad people; the problem is that the system needs to be made safer."

However, individual physicians do have some ownership in the problem. Brennan wrote of the IOM report, "… the nearly solitary focus on systems overshadowed the important and complementary role of individual physician accountability"[17]. In fact, it is estimated that approximately 1/3 of doctors will be impaired to the point of being unsafe for patient care at some point during their careers[18]. However, measuring individual physician accountability to clinical outcomes is not yet possible[19, 20].

In addition, physicians are providing care below the known state-of-the art. Just over 50% of patients in a national study by RAND in 2003 received recommended care[21]. Although Medicare spending varies by geographic location, higher levels of spending is not associated with increased quality of care[22]. And there is some evidence this level of quality of care declines throughout a physician's career.

Decline in Competence with Experience

Several sources describe a decrease in various components of professionalism over time. Medical knowledge as measured by ABIM certification is one such measure; where initial certification rates are approximately 90%, recertification rates are 86%[18]. In a recent systematic review by Harvard physician-researchers, 12 studies documented the relationship between time in practice and physician knowledge. All of them showed a consistent decline of knowledge with increasing time in practice since medical school[23]. Of note to this audience, three of these studies were of internists voluntarily taking the recertification examination.

Regarding the ultimate outcome of mortality, most studies reflect what was demonstrated in the best study: For every year since graduation from medical school, patient mortality increases by 0.5%[24]. In a recent systematic review of studies examining the association between quality of care and length of time in practice, quality was observed to decrease over time in the majority[23].

Consequences of unprofessional behavior

Unprofessional behavior may result in lost licenses, decreased quality of care, and/or malpractice lawsuits. Papadakis showed in several studies that unprofessional behavior in medical school is associated with sanctions by state boards in professional life.[25-27] Disruptive physicians may decrease quality of care by increasing errors made by nurses, students, residents, and colleagues.[18, 28, 29] A disruptive physician[18] is defined as one who

exhibits abusive behavior that "interferes with patient care or could reasonably be expected to interfere with the process of delivering quality care". Examples of such behavior are profane or disrespectful language, demeaning behavior, sexual comments or innuendos, outbursts of anger, or unethical or dishonest behavior. Physicians who have more unsolicited complaints against them for unprofessional behavior like communication, access, and availability were more likely to have lawsuits filed against them[15].

These facts taken together are disturbing: there are serious professionalism problems at every step along the march of a physician's career. These problems range across all facets of professionalism: knowledge, skills, and behavior. The consequence is poor quality of care, to which our profession has dedicated itself since its inception. These facts have just recently come to the profession's attention, and the profession is responding appropriately.

Can Professionalism be Defined?

To know whether interventions are needed and whether the interventions have resolved the above problems, one must be able to measure medical professionalism. The first step in measuring something is to define it. If something can be defined in an objective and observable manner, it can be measured. The definition of medical professionalism has been attempted by many cultures and organizations within those cultures. The uniformity of those definitions is striking.

The Hastings Center went to China in 1999 to compare the foundations of professionalism and ethics of two fundamentally different medical cultures: allopathic Hippocratic and Eastern Confucian medicine. Tom Murray commented on their similarity:

> "Every culture knows illness: and every culture makes provision for caring for people who are ill. Disease and early death disrupt the lives of individuals and families, cause physical suffering as well as great emotional pain and loss. Disease makes medicine necessary. The specific values served by medicine and the virtues cultivated in doctors flow from our shared experiences of illness, humanity, our common experiences of illness, the immense value we place on enduring human relationships, it is no surprise that we come to similar conclusions about values in medicine and virtues in physicians."[30]

Given a common set of values and virtues, they need to be articulated and linked in a practical manner to the definintion of the professional. Yet this definition has changed over time.

The Evolving Definition of the Professional

Membership in the medical profession has evolved from an entry level judgment by a preceptor after three year unstructured apprenticeship to a complex series of degrees, licenses, certification, and maintenance of certification. The future may include maintenance of licensure.

Medical Education Pre-Flexner

In early America, physicians were more likely to be barber surgeons, apothecaries, and lay practitioners because these were more likely to move to American than degreed physicians[31]. The degree was most commonly conferred after an apprenticeship, which rarely lasted more than 3 years in America (in Britain, 7 years indentured service was the rule). Another, better way to get a degree was to travel to Europe and enroll in a medical school, but not many could afford that.

Not until 1647 was any curriculum standardization attempted; dissection of a human body proclaimed by the General Court as "very necessary", but bodies were few and far between. As an illustration of the quality of instruction, Samuel Sewall recorded in his diary on 22 September 1676 that he spent the day with Dr. Brakenbury and five other observers "dissecting the middlemost of the Indian executed the day before", and that one of the group, taking the heart in hand, "affirmed it to be the stomack".

The first American hospital, Pennsylvania Hospital, was opened in 1752. The first medical school was associated with this hospital; it was called the College of Philadelphia, and was opened in 1765. Thus were born the first components of the modern medical school.

Medical Education Post-Flexner

In 1911, Abraham Flexner provided a report the effectively raised the quality and standardization of medical education for the next several decades. In response to Flexner's report, Flavel S. Thomas noted the time has passed in some schools when all that a man had to do to earn an M.D. was to "to serve as chore-boy to a physician for three years, attend two courses of lectures of four months each and pass four out of seven examinations".[31]

Since 1911, the process leading to the awarding of the degree has remained with medical schools. Most medical curricula have evolved from the Hopkins model to varying degrees.

Licensure

In the early 19th century, licensure was primitive and certification did not exist. In the nineteenth century, licenses were provided by schools, but schools proliferated and educational quality declined because of an unregulated financial incentive for the schools to give high numbers of degrees. The product was too many poorly trained physicians.

1883, Nathan Smith Davis, MD, former president of the AMA called for state boards to provide licensing examinations. States began appointing this role to the executive branch, and the governor would appoint State Medical and Examination Boards. However, some states were more proactive than others. In 1901, a letter to the editor of JAMA from a concerned consumer asked in which states people may practice medicine without being examined by the State board. The answer was published in the next issue: Arkansas, Colorado, Kentucky, Michigan, Nevada, Nebraska, Oklahoma, Rhode Island,

South Dakota, and Wyoming. In addition, being licensed in one state did not mean automatic licensure in another; there was no reciprocity.

At approximately the same time as Flexner's report, several organizations evolved whose primary mission was to improve the quality of care provided to patients. Included in these are professional societies (like the ACP), State and National Boards of Medical Examiners, and certifying boards (like the ABIM).

In 1913, the Federation of State Medical Boards (FSMB) was formed to standardize and coordinate state licensure. In 1915, National Board of Medical Examiners (NBME) formed as a result of an idea first published in a 1902 JAMA editorial to create a "voluntary, nationwide, comprehensive examination with such high standards that state licensing agencies could justifiably accept its results without the need for further examination on their parts." [32]

These two entities remained distinct until 1990, when the FSMB and NBME created the United States Medical Licensing Examinations (USMLE). The NBME hosts the first two steps of national licensure through the USMLE. Step 1 is designed to assess basic science knowledge, and Step 2, clinical science knowledge. The third and final step to licensure is sponsored by the FSMB and state boards.

A landmark shift in licensing occurred in 2004, when the USMLE added a clinical skills examination using standardized patients to license all students graduating in 2005 and thereafter. Although widely supported by the public this move was quite controversial with the American Medical Association, the American Medical Student Association, among others, primarily due to the cost[33].

The future of licensure is being heavily influenced by the need to measure professionalism: The FSMB is now talking about a "Maintenance of Licensure" program to tie in with "Maintenance of Certification" (see below). Talks are very preliminary, but all expect this to evolve over the next several years.

Certification

The growth of medical knowledge in the early twentieth century prompted the growth of medical specialties and the need to define specialty qualifications and issue credentials. Prior to the American Board of Medical Specialties (ABMS), individual physicians were the sole judge of their qualifications.[34] A key characteristic of certification is that is it voluntary; according to the president of the American Board of Internal Medicine (ABIM), "Board certification is a form of self-regulation that is funded by physicians through application fees and that remains under the control of the profession."[35]

In 1917, the first board came into existence: American Board for Ophthalmologic Examinations, soon followed by Otolaryngology, Obstetrics and Gynecology, and Dermatology and Syphilology. In 1933, the ABMS was established and by 1948, there were a total of 18 specialty boards. Today, there are 24 medical specialty boards (radiation oncology is the only prominent non-member), offering 37 general certificates

(e.g. ABIM offers certificate in general internal medicine) and 92 Subspecialty certificates (e.g. ABIM offers certificates in cardiovascular disease and transplant hepatology)[36].

In 1970, "Recertification" offered by the American Board of Family Medicine and the ABIM as a way to assure quality of care throughout a physician's career. In 1974, the ABIM was the first board to offer recertification. Through 1980, recertification remained voluntary; there were few participants when there were no repercussions for not recertifying. However, in 1982, all new certificates became time-limited in 1982 for six specialty boards. In 1990, ABIM began requiring a recertification examination at 10 years for all new certificates. Physicians with certificates conferred prior to the time limited implementation are certified indefinitely.

In 1995, ABIM was one of the first specialties to begin the "Maintenance of Certification" program – called Continuous Professional Development (CPD). This will be discussed in more detail below.

Modern Definitions of Professionalism

This history reminds us that before the 1970s cognitive knowledge was equated with professional competency. As stated previously, the advance of technology allowed physicians to better help their patients; yet this technology also distanced them from their patients. Since then, several professional societies published their own definitions of competency and professionalism.

American Board of Internal Medicine

The ABIM definition of professionalism contains three commitments and six elements. The three commitments are to the highest standards of excellence, to sustain the interests and welfare of patients, and to be responsive to the health needs of society. The six elements: altruism, accountability, excellence, duty, honor and integrity, and respect for others.

Association of American Medical Colleges

The AAMC published in 1998 the Medical Schools Objective Project[37], in which "Four Key Attributes" of professionalism were defined: altruism, knowledge, skill, and duty.

Accreditation Council for Graduate Medical Education

Working in conjunction with the ABMS, the ACGME published in 1999 an outcome project[38] that listed "six competencies" of professionalism: patient care, medical knowledge, practice-based learning, interpersonal and communication skills, professionalism, and systems-based practice. In addition, a "toolbox", which listed the use and characteristics of several methods to evaluate these competencies[39].

Institute of Medicine

In 2000, the IOM published its "To Err is Human" report, sparking the quality movement. Since then, several additional reports have been produced by the IOM,

including in 2002 the "Health Professions Education: a Bridge to Quality" report, in which were listed "five competencies for health professions education"[40].

1. *Provide patient-centered care*
2. *Work in interdisciplinary teams*
3. *Employ evidence-based practice*
4. *Apply quality improvement*
5. *Utilize informatics*

American College of Physicians/American Society of Internal Medicine

The ACP/ASIM Physician Charter[41] was published in the Feb. 5, 2002, *Annals of Internal Medicine.* This effort was co-sponsored by ABIM and the European Federation of Internal Medicine. The charter declares medical professionalism is founded on three principles: primacy of patient welfare, patient autonomy and social justice.
The charter also identifies the following 10 professional responsibilities:

1. Commitment to professional competence;
2. Commitment to honesty with patients;
3. Commitment to patient confidentiality;
4. Commitment to maintaining appropriate relations with patients;
5. Commitment to improving quality of care;
6. Commitment to improving access to care;
7. Commitment to a just distribution of finite resources;
8. Commitment to scientific knowledge;
9. Commitment to maintaining trust by managing conflicts of interest; and
10. Commitment to professional responsibilities.

The response to charter has been enthusiastic.[42] It has been published in 13 journals, translated into Italian, French, Spanish, Portuguese, German, Polish; endorsed by 90 professional organizations.

Summary: A Professionalism Construct

Given the multitude of definitions of professionalism, the consensus of the important traits is apparent[43]. A way to summarize and conceptualize the common values is presented in the following table.

Table 1: A Professionalism Construct

Cognitive	Knowledge			
	Skills	Data-gathering		
		Diagnostic		
		Management		
		Procedural		
		Computer		
Non-Cognitive	Attitudes	Communication	Language	
			Empathy	
			Integrity	
			Compassion	
		Collaboration	Responsibility	
			Respect	
			Duty	
		Continuous improvement	Recognizes limitations	
			Motivated to improve	Internal motivation
				External motivation

The literature abounds with a variety of terms addressing professionalism. The "cognitive" realm is usually used to describe those aspects of professionalism that are easier to measure, like factual knowledge measured by multiple-choice questions and manual skills measured by directly observable product. The "non-cognitive" realm generally describes more abstract aspects of professionalism that have more to do with values, attitudes, and emotions. The evaluation of these abstract concepts is difficult to imagine, which poses a problem to teaching these concepts.

Bioethical debate: Values vs. Behaviors

Over the past several years, several of the prominent thinkers in this field have begun to recognize the limitations of expressing professionalism as a collection of values. They suggest that behaviors be the focus of attention. They argue that values are immeasurable, are not part of the common experience, and are threatening in that they imply unsuitability for the profession[44, 45]. Further, the current definitions of professionalism using values are too abstract, they describe people rather than behavior, imply a stable state, and imply a true/false state for each person[46].

On the other hand, behaviors are observable, measurable, and remediable. Examples taken from the University of New Mexico school of medicine[14]:

Table 2: Linking Values and Behaviors

Value	Behavior
Responsibility	Follows through on tasks
	Arrives on time
Maturity	Accepts blame for failure
	Makes inappropriate demands

	Is abusive and critical during times of stress
Communication Skills	Listens well
	Hostile, derogatory, sarcastic
	Loud or disruptive
Respect	Maintains patient's confidentiality
	Is patient
	Is sensitive to patient's immediate physical and/or emotional needs
	Is not racist

The University of New Mexico implemented this in 1987 in an attempt to better measure professionalism, then reported their experience over three years. Ten students were identified; nine were placed into remedial programs and all but one successfully graduated. A similar process and experience was reported by the University of California San Fransisco[47]. In 2000, such a system was implemented at UT Southwestern.

Several authors suggest the need to recognize some "misbehavior" may be acceptable in some situations. For example, most physicians agree that honesty is inappropriate if it would cause the patient harm. For these reasons, once the values have been translated into behaviors, their context, conflict, and process of resolving conflict need to be taken into account[46].

If professional values can be translated into behaviors that are observable, this would increase the chances that professionalism could be measured and behaviors found wanting remediated.

Can Professionalism be measured?

There are those that believe professionalism cannot be measured[20]. Indeed, attempts to measure professionalism have been met with varying degrees of success. However, with the change in emphasis from values to behaviors, there is a growing body of literature on the measurement of professionalism.[38, 39, 45, 48, 49]

How Effective are Modern Measures of Professionalism?

Attempts to measure professionalism are apparent from admission into the profession through medical school and residency to licensing, certification, and now maintenance of certification.

For Admission into the Profession

The University of Michigan school of medicine studied the correlation between information from the application to medical school and subsequent ratings of professionalism for one class of students. This included careful reading of essays for themes of professionalism as well as history of volunteering, personal experience with illness, and expressed desire to work with underserved populations. None were found to predict professional behavior[43].

A study of the Minnesota Multiphasic Personality Inventory (MMPI) in graduate medical education makes several important points about testing for qualities prior to admission into the profession. The MMPI was developed in the 1930s by a psychiatrist and psychologist from the University of Minnesota. It has ten clinical scales: Scale 0 measures social introversion; Scale 1, hypochondriasis; Scale 2, depression; and Scale 3, hysteria. Scale 4 is the psychopathic deviate scale. Scale 5 measures masculinity-femininity; Scale 6, paranoia; Scale 7 psychasthenia; Scale 8, schizophrenia; and Scale 9, hypomania.

In this study, 40 male psychiatry residents took an MMPI (and three other tests) as an intern and were followed for 13-17 years. 2 had their licenses revoked due to repeated sexual boundary violations with patients.

These 2 were identified by the MMPI as having significant character pathology (called the "psychopathic deviancy" scale). "Individuals who score high on this scale typically have difficulty incorporating the values and standards of society and are likely to engage in lying, cheating, stealing, and sexual acting out. They also show poor judgment, take risks, and typically do not learn from past experiences."[50]

Interestingly, these residents were not identified as having problems by traditional measures: interviews, assessments by supervisors, or academic scores. These were well above average, if not excellent. The excitement generated by this realization is dampened by the fact that nine other colleagues also scored at the same level as the two who lost their license. That is, personality traits predict behavior but are very non-specific[50].

This study provides two lessons for those interested in measuring professionalism as a threshold for admission into the profession: 1) the wisdom of denying careers to all eleven residents based on this test to protect the public from two is dubious and 2) 9 of these residents have seemed to overcome their character flaws. Although the literature shows the success of counseling for character flaws like this is low, this study suggests that students and physicians with these flaws can learn proper behavior[45].

During Medical Education

The best compilation of the current state of the measurement of professionalism is presented by two groups. One group focused on graduate medical education. A second group, comprised of faculty and researchers from the medical academic community, focused on a broader population of medical students, residents, fellows, and physicians.

In 2000, the Accreditation Council for Graduate Medical Education (ACGME) performed and publicized on the web an exhaustive review of the methods available to measure the six competencies described in the Outcome Project[38, 39, 51]. Although this project focused on literature evaluating the use of these methods in graduate medical education, the results are probably applicable to medical students as well as practicing physicians.

The ACGME provides several resources for choosing which instrument is best for measuring which aspect of professionalism[39, 51]. For example, the ACGME suggests respectfulness is best measured with patient surveys or standardized patients. Teamwork is best measured with a 360 degree global rating. In all, the six competencies are broken down into 25 required skills, and the toolkit contains thirteen different evaluation methods. Each skill may be best evaluated by a different method.

In 2006, David Thomas Stern from the Univeristy of Michigan edited a compendium of chapters written by colleagues from around the country who have been studying the measurement of professionalism for years. It was published as "Measuring Medical Professionalism"[49]. This is currently the most exhaustive and up-to-date survey of methods available for assessing professionalism from medical school to retirement.

A brief summary of these two excellent and exhaustive sources of measurement of professionalism is presented in the following table and subsequent text. More detail will be provided for the most commonly used and most promising measures.

Table 3: Summary of Measures of Professionalism

Evaluation	Validity	Reliability	Practicality	Best Use
MCQ	Excellent	Excellent	Excellent	Cognitive
OSCE	Excellent	Excellent	Fair	Both
Simulations/Models	Excellent	Excellent	Fair	Skills
Attending Evaluations	Good	Fair	Excellent	Both
Peer Evaluations	Good	Fair	Good	non-cog
Patient Evaluations	Excellent	Good	Good	non-cog
360° Evaluations	Excellent	Excellent	Fair	non-cog
Oral Examination	Good	Fair	Good	Both
Procedure, Op, Case Logs	Fair	Fair	Good	Skills
Portfolios	Fair	Fair	Fair	Both

MCQ = multiple choice questions; OSCE = objective structured clinical examination; Op = operation

Multiple choice questions (MCQs)

MCQs are by far the most valid and reliable measures of the cognitive aspects of professionalism. That is, medical schools, licensing bodies, and certifying boards are quite comfortable in their abilities to accurately assess factual knowledge using written, multiple choice testing. This has led to their widespread, if not universal use. However, MCQs do not test well clinical judgment or the non-cognitive aspects of professionalism.

Objective Structured Clinical Examination (OSCE)

An OSCE is a clinical test that uses standardized patients (SPs) who have been trained by physicians to behave like real patients, yet evaluate a student or physician using a checklist of desired behaviors and skills. The OSCE has been suggested by experts in the field as one of the most promising methods of measuring professionalism[45].

It has been proved in a multitude of studies to be a reliable and valid tool for the measurement of clinical and communication skills[45, 46]. SPs are accurate raters: they demonstrate 82% agreement with the ratings of physicians[52]. They are 91% accurate when tested with standardized examinees[53]. Some students dislike the artificial nature of the examination, but SPs have been shown to be realistic, in that when they are sent unannounced to practicing physician's offices, they are only detected 10% of the time. One third of these detections occurred because the SPs did not fit physician's practice[52]. They have been shown to accurately portray cases at a 93.4% level[54]. Students behave the same whether seeing SPs or regular patients in terms of empathy[55] and numbers of questions asked[56].

OSCEs have also been explored as a method to evaluate the more complex behaviors inherent in the clinical encounter. OSCEs underemphasize important aspects of professionalism, like integration of knowledge and skills, information management, and commitment to improving health systems[48]. In addition, they have been used to assess ethical behavior. Given the complexity of ethical situations, experts estimate 41 stations would be required for adequate reliability[46].

One dramatic example from the literature demonstrates how OSCEs can be useful in discovering lapses in professionalism that might otherwise be missed. This example comes from a psychiatric OSCE:

> "… we detected a number of worrisome incidents not previously reported in the OSCE literature. Specifically, we saw several cases in which students demonstrated violations of professional conduct in the context of a station performance (e.g., the assaultive restraint of an SP portraying schizophrenia, inappropriate responses to the scripted sexual advances of an SP portraying mania)."[57]

The USMLE Step 2 uses an OSCE to assess student data gathering, language, and communication as part of the licensing process. They have found it to be very reliable.

Simulators and models have not been evaluated extensively in the literature, but when they have, they have been used to test skills. For this purpose, they have shown psychometric qualities similar to standardized patients.

The disadvantage of OSCEs, simulated patients, and mechanical models are the costs. OSCEs are challenging to design, time-consuming to administer, and costly to implement. Standardized patients must be trained, checked out, and paid. Good mechanical models are very costly.

360 Degree Evaluations

Evaluations may be performed by attending physicians, peers, patients, nurses, consultants, referring physicians, and other ancillary staff. The use of all these sources is employed in a 360 degree evaluation. This type of evaluation has been championed in the business world, where its psychometric qualities (validity and reliability) have been

documented to be quite impressive. Its psychometric qualities have yet to be studied in the medical world, but there are few reasons to believe the psychometric qualities would be much different.

Peer evaluation

Although evaluation by peers is faced with problems of standardization and less expertise than supervising faculty, there are many benefits offered only by this form of evaluation[58]. At face value, peers have more contact with each other in the context of patient care than any other group (except perhaps nurses).

In general, students are supportive of peer evaluation, in the appropriate context. Sixty-six percent of the students surveyed at one school agreed that there should be peer assessment of professionalism as long as the assessment was reported to an impartial counselor, was done anonymously, and resulted in corrective instruction, rather than a lower grade.[59] Indeed, if peer evaluations are part of a high-stakes outcome (like a grade, degree, license, or certificate) rather than for educational purposes, they become much less useful, and probably not worth the effort of collecting them.

Several types of peer evaluation have been evaluated in the literature. Ratings are most common and best studied of the types of peer evaluation. For example, peers nominate classmates for awards or special recognition, and typically evaluated on a Likert scale. Ratings are good for identifying particularly good or bad performers, but are less useful for uniform, individual evaluation of an entire class of students.

Another type of peer evaluation is ranking. In this method, peers are ranked regarding a global characteristic, like compassion, empathy, or even professionalism in general. Rankings are the best way to obtain a reliable evaluation on every member of the class.

Research shows that peer evaluation is measuring something in a reliable and valid way. Since it does not correlate well with MCQ scores, it is different than knowledge or skill. Peer evaluation requires six to eleven evaluations per individual to be reproducible.

Peer evaluations have several limitations. Both ratings and rankings are evaluations based on global (rather than specific) characteristics and based on values rather than behaviors. This limits their usefulness as a teaching tool. In addition, peer evaluations can be affected by group dynamics (e.g. attending style). They are influenced by consequence. That is, higher stakes ratings become less reliable, more homogeneously high.

Supervising Physician Evaluations

This is the most commonly used type of evaluation for professionalism. The evaluation is done by faculty of students, residents, or fellows, usually on ward or clinic rotations. There are many problems with this type of evaluation, especially for professionalism. The evaluations are relatively unreliable, and supervisors are reluctant to provide negative feedback.

The literature shows that rating forms of resident-patient interactions: reliability would require many evaluations. For example, 20-50 attending physician evaluations per resident are required, 50 patient ratings per resident, 5-10 program supervisors, and 10-20 nurses[60]. Interestingly, nurses' evaluations correlate best with patient evaluations, followed by attending physicians, then program supervisors.

In addition to the need for multiple evaluations, individual faculty members are reluctant to provide negative feedback and evaluation[14]. One study used a multi-method, qualitative case study approach to studying four teams over 40 hours each over two weeks. Semi-structured interviews, observation by a physician, structured cognitive task with thinking aloud. The four attending physicians were selected in order to achieve a balance of gender, age, subspecialty, and reputed teaching style. Each had a minimum of 8 years teaching experience.

Attending physicians identified three main "red flags": cutting corners (e.g. avoiding admission, early discharge, reluctance to pursue clinically appropriate steps), disrespect (e.g. disparaging remarks about patients), and outright hostility, malice, or rudeness. Attending physicians appeared more reluctant to intervene in practice than they described themselves to be in the think aloud exercise:

"Dr. Arthur, for example, took umbrage at the use of the term 'shooter' in the think aloud exercise, and said he would stop rounds and 'lay down the law' if ever an intern called a patient a 'dirtball'. "But he showed no discernible reaction when team members regularly referred to injection drug users as 'shooters,' or when his resident described obese patients with pulmonary disease as 'slug-like'."

Attending physicians "did not feel supported or rewarded by their institutions for trying to teach attitudes. They believed that doing so not only would not advance their academic promotion or success, but might even open them to personal and legal risk."[61]

Brief digression: Do the courts support evaluation of professionalism?

This question is paramount to the proper evaluation of professionalism, so it warrants a digression. One of the many obstacles to evaluating professionalism is a faculty concern about legal repercussions resulting from a negative evaluation[62]. If the question is "Can I/the University get sued for a negative evaluation that results in dismissal", the answer is "yes". This is part of living in a free and unfortunately litigious society. However, UT Southwestern faculty should know that the both our campus and the University of Texas Attorney General's office provide legal assistance for state employees who are sued as a result of their duties. In addition, if the court rules against the faculty member, the UT System indemnifies state employees for up to $100,000.

The real question should be: if sued, who shall prevail? The answers to this question are rooted in three legal principles: procedural due process, substantive due process, and defamation.

For the purposes of this discussion, the focus shall be on the worst case scenario: a student has been dismissed because of unprofessional behavior. The layperson's interpretation of these legal principles can be applied to this scenario, resulting in legal questions more specific than simply "who shall prevail". For example:

1. Is the process so complex that dismissal will be overruled on a technicality? (procedural due process)
2. Will the courts find observations of unprofessional behavior too subjective to result in the termination of a student's career? (substantive due process)
3. Will the courts find negative, non-cognitive observations are libelous and slanderous? (defamation)

In short, the answers to these questions are "no". This section will provide the evidence to support these answers for the first two.

Please note this discussion is best applied to undergraduate medical education at State supported schools. Legal decisions at private schools, residencies, and fellowships are guided by contract law, while public schools are held to both contract and constitutional law. Contracts are variable and unique to each institution, but are also generally based on rights guaranteed by the constitution. Although there may be many parallel general concepts between contracts and constitutional law, the nuances discussed here are most applicable to State supported schools.

Due Process

Most suits brought by medical students as a result of their dismissal are based on 14th amendment to the U. S. Constitution's "due process clause". This clause states, "...nor shall any State deprive any person of life, liberty, or property, without due process of law". The courts have been divided on whether a student's dismissal deprives them of liberty (e.g. freedom to pursue a career) or property (e.g. time and money invested in pursuing this career). Therefore, it is wisest to assume students have this liberty or property interest until the Supreme Court clarifies this point.

Assuming students have this interest, the 14th amendment states citizens may not be deprived of liberty or property without "due process". The Supreme Court provides some guidance here. In Horowitz[63], it stated there are two types of "due process": procedural and substantive. As this case provides a good example to follow, it will be reviewed in more detail below. Procedural due process defines *how* students are to be dismissed. Substantive due process defines *why* students are to be dismissed.

Procedural Due Process

The potential list of requirements for due process is long; a partial list is provided in the following Table.

Table 4: Partial List of Potential Rights in Procedural Due Process

(1) Right to notice of deficiency/violation
(2) Right to hearing
(3) Right to pre-hearing notice of evidence

19

(4) Right to preservation of pre-hearing status	
(5) Right to impartial hearing board	
(6) Right to counsel	
(7) Right to confront witnesses	
(8) Right to invoke privilege against self incrimination	
(9) Right to obtain a record of the proceedings	

The appropriate level of procedural requirements varies according to the facts of the case and the severity of the penalty. That is, criminal cases involving jail time or worse require more procedural protections than scholastic dismissal cases. The procedural protections required in university cases dependent on whether the dismissal is due to misconduct or academic reasons.

In general, misconduct is defined as violation of University rules or regulations. Examples include cheating, plagiarism, and criminal conduct like stealing, drug abuse, drunkenness while on duty, and falsifying (academic) records. As defined in Horowitz[63], academic dismissals are those based on cognitive as well as non-cognitive (clinical) performance; examples include hygiene and timeliness.

In Dixon[64], a dismissal due to misconduct requires the university to provide first three (notice [should be written], hearing, evidence). Dismissals for misconduct are handled by administration, and a discussion of this is beyond scope of this talk. The disciplinary policy is outlined in the UT Southwestern catalog[65] and is provided in detail in series 50101 of the Regents' Rules of the University of Texas System Board of Regents[66].

One of the main messages from the Supreme Court in Horowitz[63] is if a dismissal is due to a violation of academic standards, only minimal notice is required. That is, students **must** be notified that passing courses is requirement for graduation (catalog will suffice), and students **must** be notified of their deficiency (students **should** also be given time to improve).

Substantive Due Process

Per the Connelly case[67], substantive due process requires dismissal **must not** be in bad faith, arbitrary, or capricious. That is, dismissal must be based on judgments that are fair and equitable. The courts have long ago[68] defined bad faith: "breach of faith, willful failure to respond to plain, well-understood statutory or contractual obligations". Courts [69]have also defined arbitrary and capricious: "willful and unreasoning action, without consideration and in disregard of facts or circumstances".

Several cases have applied these definitions fairly clearly in the setting of academic dismissals. The U. S. Supreme Court declared in Ewing[70] that courts should show "great respect for faculty's professional judgment". This "great respect" will be further defined below in a more detailed review of this case, but a meaningful reflection of it is found in the fact that the *burden of proof is on the student* to prove a dismissal was in bad faith, arbitrary, or capricious[67].

Other courts[67] have declared, "… The rule of judicial nonintervention in scholastic affairs is particularly applicable in the case of a medical school. A medical school must be the judge of the qualifications of its students to be granted a degree; Courts are not supposed to be learned in medicine and are not qualified to pass opinions as to the attainments of a student in medicine".

Despite this latitude provide by the courts, schools should follow policies promulgated by University, as this is the purview of contract law. Schools can create rights for students that have no constitutional basis by what is published by the University in policies and catalog. Once a policy is written, it must be followed exactly, or the university can be found wanting at the court house.

Illustrative Due Process Cases

Prior to 1961, there was a tradition of judicial non-interference with academic issues. Students were dismissed for a variety of reasons, and they could find no recourse in the courts. Over the next several decades, things changed: admission into medical school and residency became much more difficult, student debt rose, student access to and rights regarding their educational records increased, and society became generally more litigious[62].

In Dixon[64], one case changed the rule of judicial non-interference with academic issues forever. Alabama State College expelled five students because of their participation in a civil rights demonstration. This expulsion was done without notice, hearing, or appeal. In short, they were expelled without procedural due process. Dixon, one of the students, brought suit against the Alabama Board of Education and ultimately the 5th Circuit Court of Appeals brought constitutional protections to education. They established the procedural due process requirements of notice and hearing before any student could be dismissed for misconduct.

This case opened the court doors to dismissed students, and the next several years were marked by inconsistent decisions as the courts tried to apply constitutional protections to complex academic cases. It was unclear whether dismissal cases required all the proceedings afforded criminal cases. Were students required to have legal representation? Were all student proceedings to be recorded and transcribed? It wasn't until eighteen years later that the Supreme Court provided guidance.

The Horowitz Case

The first significant guidance on due process in dismissal cases was provided in the Board of Curators of the University of Missouri v. Horowitz[63]. In August 1971, Charlotte Horowitz was admitted with advanced standing to year 5 (of 6) at the University of Missouri, Kansas City medical school. One year later, in the Summer of 1972, an annual review of all students revealed her attending physician in the pediatrics course "criticized her lack of patient rapport and expertise in coming to the fundamentals of the clinical problem, erratic attendance, and poor personal hygiene."

Horowitz was advanced to year 6 on probation. She was judged as "unsatisfactory" on a subsequent rotation, and a committee recommended her dismissal. As part of her appeal of dismissal, a committee required review of Horowitz's work through oral and practical examinations by seven independent clinicians: 2 said she should be allowed to graduate, 2 said she should be dismissed, three said she should not be allowed to graduate yet, and continue on probationary status.

In the Summer of 1973, after several more low grades, a panel of reviewing clinicians recommended her dismissal to the dean, which was upheld through appeals to dean and provost. Horowitz filed suit, claiming the dismissal deprived her of freedom to continue in school and become a doctor (liberty interest), and therefore the 14th amendment required due process. Further, she claimed *procedural* due process was violated, since hygiene and timeliness are "conduct" issues, rather than "academic", therefore, hearing was required.

The District court ruled in favor of the school. In 1976, the 8th circuit court of appeals ruled in favor of Horowitz.

In 1978, the U. S. Supreme Court ruled a hearing was not required:

> "Academic evaluations of a student, in contrast to disciplinary determinations, bear little resemblance to the judicial and administrative fact-finding proceedings to which we have traditionally attached a full-hearing requirement. ... We decline to ignore the historic judgment of educators and thereby formalize the academic dismissal process by requiring a hearing."

> "A school is an academic institution, not a courtroom or administrative hearing room."

> The court "declined to further enlarge the judicial presence in the academic community and thereby risk deterioration of many beneficial aspects of the faculty-student relationship."

Further, it ruled that professionalism is part of an academic performance:

> "The decision to dismiss [Horowitz] ... rested on the academic judgment of school officials that she did not have the necessary clinical ability to perform toward that goal. Such a judgment is by its nature more subjective and evaluative than the typical factual questions presented in the average disciplinary decision."

> "Personal hygiene and timeliness may be as important factors in a school's determination of whether a student will make a good medical doctor as the student's ability to take a case history or diagnose and illness."

In summary, we can take the following points from Horowitz. In this case, the Supreme Court distinguished "Academic" from "Disciplinary" dismissal. In academic dismissal,

the only procedural due process required was "notice". The Court further defined "notice" as rules in catalog. Legal analysts since then state informing the student of his/her deficiency and providing time to improve solidifies the "academic" nature of these actions and improves school's stance in court.

Finally, and most germane to this discussion of professionalism, the Supreme Court defined non-cognitive, clinical performance (i.e., in this case timeliness and personal hygiene) as "Academic".

The Ewing Case

Although Horowitz did much to clarify what *procedural* due process was required in dismissal cases, guidance was still needed regarding what *substantive* due process was required. In other words, in what circumstances could courts overrule a school's dismissal of a student?

This guidance came again from the U. S. Supreme Court in Ewing v. the University of Michigan School of Medicine[70]. In August of 1975, Scott Ewing entered U. Michigan's 6 year medical school. After 6 years, only 4 years of class work was completed, marked by personal problems, inadequate preparation, poor grades, incompletes, withdrawals. At the time, all U. M. students were required to pass NBME step 1 examination prior to clinical rotations. Passing was defined as a score of 345; Ewing made a 235, the lowest score ever recorded at U. M.

Ewing was expelled, appealed citing several purportedly extenuating circumstances, but his expulsion was upheld. Ewing filed suite, claiming his dismissal deprived him of property (time and money invested in degree – i.e. property interest), and therefore the 14th amendment required due process. Further, he claimed his *substantive* due process was violated, because dismissal was arbitrary and capricious.

In 1983, the District court ruled in favor of U. M. In 1984, the 6th Circuit Court of Appeals ruled in favor of student. A U. M. pamphlet was found after the district court ruling that stated, "a qualified student would be given a second chance to take the NBME". Therefore, 6th Circuit Court found that Ewing should be allowed to take NBME exam second time. The dismissal was "arbitrary and capricious" because Ewing was treated differently than other students.

In 1985, the Supreme Court overruled the 6th Circuit decision. They declared:

> "When judges are asked to review the substance of a genuinely academic decision, such as this one, they should show great respect for the faculty's professional judgment. Plainly they may not override it unless it is such a substantial departure from accepted academic norms as to demonstrate that the person or committee responsible did not actually exercise professional judgment"

In summary, the following points can be taken from Ewing. This case further defined "arbitrary and capricious" as when professional judgment *was not used*. Essentially, the Supreme Court gave schools the ability to make the wrong decision for the right reasons.

Substantive due process is honored when an individualized review of entire record is basis of the decision, when there is no substantial departure from accepted academic norms, and the person or committee making the decision exercised professional judgment.

In summary of this section on the current measurement of professionalism, the battery of instruments available to measure professionalism is growing in volume and quality. However, a recent review found that very few of these instruments yet have acceptable psychometric properties to warrant use in high stakes decisions[71].

How Can the Measurement of Professionalism be Improved?

Several suggestions for improving the measurement of professionalism are apparent in the literature. A change in thinking of professionalism as a collection of values to a collection of behaviors has already been discussed. Prominent thinkers in the field suggest an additional change in philosophy is needed.

Reductionist vs. Expansionist perspective

Russell Lincoln Ackoff[72] is viewed as the father of Systems Theory, which is now embraced by the "interdisciplines" of cybernetics, decision sciences, management sciences, computer science, ecology, and systems engineering (e.g. NASA and the Jet Propulsion Laboratory). Systems theory is credited with W. Edwards Deming's successful transformation of the Japanese auto industry in the 1950s that contributed to their economic boom. Systems theory distinguishes analytic thinking and its resulting reductionism from synthetic thinking and its expansionism.

The successes of the Renaissance, the Enlightenment, and the Industrial Revolution are largely a product of analytic thought, where answers to questions were discovered by analysis: by taking them apart, breaking them down, and viewing them at higher and higher magnification. The Systems approach is to identify the whole of which the object of the question is a part, explain the behavior or properties of the containing whole, and then explain the behavior or properties of the object to be explained in terms of its role(s) or functions(s) within its containing whole[73]. An example of the benefit of a systems approach is provided by Dr. Ackoff:

> "An all-star football team is seldom as good as the best team in the set from which all the players are drawn. But, you might say, if the all-stars would play together for a while, they might become the best team. Yes, but when they do, some, if not most, would no longer be selected as all-stars."

However, there is no "right" or "wrong" way to approach questions. In Dr. Ackoff's words,

> "Neither way of thinking negates the value of the other but by synthetic thinking; we can gain understanding of individual and collective human behavior that cannot be obtained by analysis alone."

Ackoff wrote much about systems theory in education, much of which can be applied to professionalism. We have seen in the above definitions that professionalism has been defined is a multi-faceted, complex entity within the health care system. Many feel the assessment of professionalism would be best conducted in an expansionist manner. That is, "identify the universe of elements of professionalism" that contribute to the function of the health care system[74].

Such an assessment might take the form of a multi-score profile constructed by gathering outcome date from various interactions of the professional within the health care system. 360 degree evaluations is one example provided by the business world. Other examples are explored below. First, the desired type of evaluation must be addressed.

Formative vs. Summative Evaluation

In general, there are two ways in which evaluations can be used. The distinction between these two types of evaluation can be summarized by the words of Robert Stakes: "When the cook tastes the soup, that's formative evaluation; when the guests taste the soup, that's summative evaluation". Medical educators are the cooks preparing a "soup" for the public. It is these educators' duty to the public to perform summative and their duty to their students to perform formative evaluations.

The value of formative evaluation is obvious; students receive formative feedback on their knowledge base through the seemingly endless series of multiple-choice question examinations. Students will adjust how and what they study according to this feedback.

Such feedback is also essential for non-cognitive aspects of professionalism. However, students rarely receive formative feedback on these essential non-cognitive aspects for many reasons. Supervising attending physicians have "difficulty in identifying problems, an inability to verify problems, and a fear of litigation"[14]. Even if problems are identified, it has been shown above that supervising physicians are reluctant to act upon bad behavior in a manner they declare is appropriate[61].

These difficulties encountered when trying to assess professionalism has fueled a debate in the bioethics literature.

Pioneering Applications of these Philosophies

Several groups are embracing the Systems Theory and formative approach to professionalism. Examples from across the spectrum of medical education to practice are provided below. In medical school, Hojat and colleagues at Jefferson Medical College are exploring a multi-score profile. At the other end of the spectrum, certification boards in Canada and America are aligning along these theories. These efforts are new, and studies of their efficacy are in progress. However, some data supporting their use are becoming available.

Jefferson Medical College

An interesting application of the systems theory of Ackoff has been applied by researchers at Jefferson Medical College[74]. Supported by the Stemmler Medical

Education Research Fund, they developed and verified several questionnaires developed to assess various aspects of professionalism: empathy, collaboration, and life-long learning. Students complete the forms and receive individual scores for each of these aspects. These individual scores are combined into a multi-score profiles created based on Systems Theory[48].

Empathy

The Jefferson researchers' vision of empathy encompasses the humanistic qualities respect, compassion, honor, and integrity. The Jefferson researchers report good reliability, primarily through assessing internal consistency through Cronbach's alphas. In terms of validity, factor analysis supports the instrument measuring something. In addition, the construct validity of the instrument is supported by gender studies (where women perform better than men), the lack of correlation with cognitive measures, the correlation with clinical evaluations, and the correlation with students' choice of "people specialties" for their careers.

Examples of items from the empathy questionnaire:

> "I try to understand what is going on in my patient's mind by paying attention to their nonverbal cues and body language."

> "I try not to pay attention to my patients' emotions in history taking or in asking about their physical health."

> "Because people are different, it is difficult for me to see things from my patients' perspectives."

Others in the field believe the assessment and teaching of empathy can go beyond self-evaluation. Based on a concept first suggested in Hochschild's 1983 book, The Managed Heart[75], Eric Larson suggests empathy should be viewed as "emotional labor", which can be taught and measured much like flight attendants can be taught to be pleasant despite unreasonable requests and actors can be taught to assume any role[76]. Larson and others[77] believe a set of behavioral components of empathy and patient rapport exists and medical schools and residencies should play an active role in training and development of empathy[76].

Collaboration

The Jefferson group defined collaboration as encompassing the values of respect, accountability, and duty. This seems inherently important as a component of professionalism, and there are data from the APACHE studies suggesting collaboration improves survival in intensive care[78]. The validity of this instrument is supported by cross-cultural studies. For example, the health care environment in the United States and Israel is much more accepting of collaboration of nurses in the care of patients than the environments in Italy and Mexico. Physicians in the U.S. and Israel received higher scores than Italian and Mexican physicians.

Lifelong Learning

The Jefferson researchers defined lifelong learning as maintenance of competence, self-regulation, and participation in quality improvement activities. They then derived and validated a 19-item self-assessment tool to measure these characteristics. Incorporated in this tool are assessments of research activity, internal and external educational motivation, and computer skills.

The Jefferson group's experience with these scales is new, and they are just beginning to explore the utility of these scales in medical education. For example, they have yet to be incorporated into the routine evaluation and feedback mechanisms for all students attending Jefferson[79]. However, certifying boards in North America are exploring the use of other methods of multi-score profiling.

Alberta Licensing Board Physician Achievement Review

The College of Physicians and Surgeons of Alberta, which is required by law to review the licenses of all physicians every five years, assesses professionalism with confidential surveys of practicing physicians every five years, administers a multi-score profile that assesses professionalism. Results with standardized norms are provided to each physician[80].

Participating physicians receive a package of questionnaires and a self-evaluation form. There are questionnaires for patients, co-workers and peers that cover five attributes of the physician's performance: clinical knowledge and skills, communication skills, psychosocial management, office management and collegiality.

Once the questionnaires have been completed, they are combined into PAR profiles that are returned to the physician and reviewed by a committee of appointed physician peers. The role of the PAR profiles and committee review is primarily formative feedback. That is, if the PAR surveys flag a potential problem, the committee works with the physician from a quality-improvement perspective.

This process embraces the formative evaluation stance as well as Systems Theory. The PAR website states, "PAR information is strictly for educational purposes and cannot be used in legal or disciplinary proceedings." However, there is a clause that does allow intervention, if needed: "Only in those rare instances when it is determined that, 1. the public is at immediate risk of harm, 2. there is a serious breach of ethics, or 3. there is an unreasonable failure to comply with the requirements of the PAR process, could a physician be referred for possible formal investigation."

The PAR process is also relatively new, and measures of its success in improving quality of care are currently unavailable. In addition, the psychometric characteristics of these instruments were not readily discovered.

ABMS Maintenance of Certification (MOC)

The ABMS's evolution into the MOC system is described above, and probably represents the most advanced application of Systems Theory and a move toward formative evaluation. The MOC program is designed to be a continuous evaluation of physician's competence beyond what knowledge is tested on written examinations. Starting in January 2006, MOC options for evaluation of practice performance included peer-and-patient feedback and quality improvement modules[35].

There are four components of MOC:

1. Evidence of professional standing (an unrestricted license to practice medicine)
2. Evidence of lifelong learning (evidence of CME participation)
3. Evidence of cognitive expertise (a secure online MCQ test)
4. Evidence of evaluation of practice performance (several options exist, including patient and peer surveys, chart reviews, involvement in quality improvement activities at the home site)

At present, all specialties register a 87% certification rate. At the time of the 2002 survey documenting this rate, not all specialties participated in MOC. ABIM has the most experience with MOC and recently published its data[81]. 77% of general internists and 60% of subspecialists participate in MOC; the primary reason stated for not participating is the time-consuming nature of the MOC process. Most subspecialists that do not participate in the general internal medicine MOC state it was no longer relevant to their subspecialty practice.

The lower participation rates in internal medicine than other specialties highlight the main disadvantage of the MOC process: it discourages participation. Most that do participate in MOC state it is a requirement of one or more of their employers, but the main reason they do it is for positive professional reasons, like maintaining or improving one's professional image, updating knowledge, and maintaining or improving quality patient care and safety[81].

Although MOC is relatively new, some data to assess its efficacy are becoming available. 356 physicians that participated in the MOC between 1999 and 2002 were surveyed about their experience. The patient, peer and self evaluations module was the focus; these physicians found the experience to be a valuable learning experience and the results could be used to improve their practices"[82]. The module was also found to be valid and reliable using standard psychometric measures.

One component of the MOC process, written assessments of knowledge, is associated with outcomes more concrete than survey results. Board certification is associated with higher scores on a standardized knowledge test as well as higher ratings by peers[83]. Regarding patient outcomes, the majority of the best studies demonstrate a positive relationship between certification and patient outcomes, but the results are not unanimous[17, 83, 84]. Norcini showed in one of the best studies of the issue that physician

board certification is associated with better outcomes in patients with acute myocardial infarction (15% decrease in mortality)[24].

Given the positive response and outcomes of these results as well as the rising expectations of the public driven by the quality movement, each of the 24 member boards of ABMS agreed in 2002 to implement "Maintenance of Certification" by 2010[85].

The MOC program is becoming less voluntary. Two recent articles in JAMA relate how certification is becoming increasingly linked to health plan credentialing and hospital privileging. In the first article, researchers explored the relationship between certification and health plan credentialing: 10% require initial certification, 41% require eventual certification, and 41% require recertification (if certificate is time limited)[86].

The second article focused on hospital privileging of pediatricians. 22% require initial certification, 70% require eventual certification, and 45% require recertification (if certificate is time limited).[87]

3% of "grandfathered" physicians recertify[88]. There were two recent editorials in NEJM from such physicians, making a case for all physicians to recertify[89, 90]. Troy Brennan, a professor of medicine at the Brigham and Women's Hospital and Harvard Medical School writes,

> "So I wouldn't have attempted recertification were it not for a nagging sense of hypocrisy and my unease with some relevant clinical issues. I saw a lack of integrity in belonging to a profession that forced younger physicians to become recertified while I declined to do so myself. My discomfort was compounded by a letter someone wrote to the American Board of Internal Medicine (ABIM) complaining about the recertifying exam and pointing out that I, the ABIM chairman at the time, had never been recertified."

Dr. Brennan goes on to describe his experience as very positive. The preparation for the examination by using online modules as part of the "evidence for life long learning" component was relevant to his practice of medicine. The second editorial describes the experience with the practice improvement module as being equally positive. Both agree that the MOC is time consuming, but both conclude the increase in confidence and quality of care as a result of participating makes it worth the investment.

Given the low failure rates for certification, the primary role is formative evaluation. Nonetheless, something must be done with those that fail. Several authors call for an expansion of the education and remediation programs currently available; today only ten programs exist nationwide[18].

Conclusion

Possibility of reliably and accurately measuring most of the non-cognitive aspects of professionalism is not yet a reality[71]. Great progress is being made as techniques are researched and refined. A change to a behavior- rather than value-based definition of

professionalism will help tremendously. A reductionist search for the perfect single tool to measure professionalism is unrealistic; an expansionist, multi-score profile offers the most promise. Finally, too little attention is paid to teaching professionalism; this must be addressed before it is measured in a high-stakes situation like grades, degrees, licenses, or certification.

Medical school faculty are an important part of the teaching and evaluation of professionalism. In addition to the evaluations being primarily formative and constructive, the evaluations should have high volume and frequency. Evaluations should be based on behaviors rather than values. When providing summative feedback, even when it may result in the ending of a career, evaluations should be unhindered by legal concerns.

Some of the foremost thinkers in this field of the metrics of professionalism write,

> "There appears to be an unrealistic expectation that students will arrive at medical school lacking in knowledge and skills, but with a full complement of appropriate behaviors that require no further attention. However, all students are vulnerable to lapses in professional behavior and can benefit from explicit, systematic attention in this domain. The focus of medical education in the past century was on knowledge and skills. For the future of medicine, attention to the teaching and evaluation of professionalism is vital.[46]"

Finally, and perhaps most important, faculty must be a good role models. This means demonstrating the highest standards in professional conduct in the care of patients, in interactions with colleagues and ancillary staff, and as an educator. This also means tending to the quality of care provided by ourselves through participating in MOC programs whether or not we have a time-limited certificate.

There is good news from the cynicism literature. First, it appears medical student cynicism is temporary. Testerman found significantly lower levels of cynicism and hostility and higher levels of optimism as one moves from student to resident to faculty members[91].

Second, the cynicism literature highlights importance of attending physician role modeling in student cynicism. Wear discovered that the behavior and misbehavior of attending physicians had a powerful effect on students. She writes,

> "It appears that students hold attending physicians to a different, perhaps higher, set of expectations. It may be that medical students identify with the 'hellish' existence of residents, forgive them for their transgressions, and join in as a form of empathy for their suffering. Equally plausible is that medical students recognize that the ultimate responsibility for the care of patients and for the outcomes of that care lies with the attending. Students may surmise that the patient, in the final analysis, looks to that attending as the 'real doctor,' and thus is owed a higher level of respect in that relationship."[8]

Being a member of the ranks of physicians is an honor and privilege. Being an internal medicine physician carries this to an even higher level, as internal medicine has lead the journey toward improving quality of care and professionalism. Internal medicine was the first to offer recertification and was the very proactive in the development of a national, then international definition of professionalism. The need to improve professionalism is great, and the chosen course of action offers great hope for meeting this need.

References

1.   ABIM. *Project Professionalism*. Philadelphia: American Board of Internal Medicine; 1995.

2.   Sierles F HI, Circle S. Cheating in medical school. *Journal of Medical Education*. Feb 1980;55(2):124-125.

3.   Becker H GB, Hughes E, Strauss A. *Boys in White*. Chicago: University of Chicago Press; 1961.

4.   Eron LD. The effect of medical education on attitude: A follow-up study. *Journal of Medical Education*. 1955;33(2):25-33.

5.   Wolf TM BP, Faucett JM, Randall HM. A retrospective study of attitude change during medical education. *Medical Education*. Jan 1989;23(1):19-23.

6.   Sheehan KH SD, White K, Leibowitz A, Baldwin DC Jr. A pilot study of medical student 'abuse'. Student perceptions of mistreatment and misconduct in medical school.[see comment]. *JAMA*. Jan 26 1990;263(4):533-537.

7.   Feudtner C CD, Christakis NA. Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Academic Medicine*. Aug 1994;69(8):670-679.

8.   Wear D AJ, Varley JD, Zarconi J. Making fun of patients: medical students' perceptions and use of derogatory and cynical humor in clinical settings.[see comment]. *Academic Medicine*. May 2006;81(5):454-462.

9.   Hafferty FW. Beyond curriculum reform: confronting medicine's hidden curriculum. *Academic Medicine*. Apr 1998;73(4):403-407.

10.   Ginsburg S RG, Lingard L. The disavowed curriculum: understanding student's reasoning in professionally challenging situations. *Journal of General Internal Medicine*. Dec 2003;18(12):1015-1022.

11.   Stewart JB. *Blind Eye*. New York: Simon & Schuster, Inc.; 1999.

12.   Geringer J. Michael Swango: Doctor of Death. http://www.courttv.com/onair/shows/mugshots/indepth/swango.html. Accessed 21 June, 2006.

13.   Levy B. 1996:Personal Communication to Wagner JM,.

14.   Phelan S OS, Galey WR. Evaluation of the noncognitive professional traits of medical students. *Academic Medicine*. Oct 1993;68(10):799-803.

15.   Hickson GB FC, Pichert JW, Miller CS, Gauld-Jaeger J, Bost P. Patient complaints and malpractice risk.[see comment]. *JAMA*. Jun 12 2002;287(22):2951-2957.

16.   Kohn LT CJ, Donaldson MS, ed. *To Err is Human: Building a Safer Health Care System*. Washington D.C.: National Academy Press; 2000. Institute of Medicine.

17. Brennan TA HR, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification status in the quality movement.[see comment]. *JAMA*. Sep 1 2004;292(9):1038-1043.

18. Leape LL FJ. Problem doctors: is there a system-level solution? *Annals of Internal Medicine*. Jan 17 2006;144(2):107-115.

19. Hofer TP HR, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease.[see comment]. *JAMA*. Jun 9 1999;281(22):2098-2105.

20. Landon BE NS, Blumenthal D, Daley J. Physician clinical performance assessment: prospects and barriers.[see comment]. *JAMA*. Sep 3 2003;290(9):1183-1189.

21. McGlynn EA AS, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States.[see comment]. *New England Journal of Medicine*. Jun 26 2003;348(26):2635-2645.

22. Fisher ES WD, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care.[see comment][summary for patients in Ann Intern Med. 2003 Feb 18;138(4):I36; PMID: 12585853]. *Annals of Internal Medicine*. Feb 18 2003;138(4):273-287.

23. Choudhry NK FR, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care.[see comment][summary for patients in Ann Intern Med. 2005 Feb 15;142(4):I54; PMID: 15710955]. *Annals of Internal Medicine*. Feb 15 2005;142(4):260-273.

24. Norcini JJ KH, Lipner RS. Certification and specialization: do they matter in the outcome of acute myocardial infarction? *Academic Medicine*. Dec 2000;75(12):1193-1198.

25. Papadakis MA HC, Teherani A, Kohatsu ND. Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. *Academic Medicine*. Mar 2004;79(3):244-249.

26. Papadakis MA TA, Banach MA, Knettler TR, Rattner SL, Stern DT, Veloski JJ, Hodgson CS. Disciplinary action by medical boards and prior behavior in medical school.[see comment]. *New England Journal of Medicine*. Dec 22 2005;353(25):2673-2682.

27. Teherani A HC, Banach M, Papadakis MA. Domains of unprofessional behavior during medical school associated with future disciplinary action by a state medical board. *Academic Medicine*. Oct 2005;80(10 Suppl):S17-20.

28. Benzer DG MM. The disruptive-abusive physician: a new look at an old problem. *Wisconsin Medical Journal*. 1995;94(8):455-460.

29. Diaz AL MJ. A definition and description of nurse abuse. *Western Journal of Nursing Research*. Feb 1991;13(1):97-109.

30. Arnold EL SD. What Is Medical Professionalism? In: Stern DT, ed. *Measuring Medical Professionalism*. New York: Oxford University Press, Inc.; 2006.

31. O'Malley CD. *The History of Medical Education*. Vol 12. Los Angeles: University of California Press; 1970.

32. Hubbard JP LE. *The National Board of Medical Examiners: the First Seventy Years*. First ed. Philadelphia: National Board of Medical Examiners; 1985.

33. Papadakis MA. The Step 2 clinical-skills examination.[see comment]. *New England Journal of Medicine*. Apr 22 2004;350(17):1703-1705.

34. ABMS. History of the ABMS. http://www.abms.org/history.asp#. Accessed 24 June, 2006.

35. Steinbrook R. Renewing board certification. *New England Journal of Medicine*. Nov 10 2005;353(19):1994-1997.

36. ABMS. American Board of Medical Specialties; 2005.

37. AAMC. *Report I: Learning Objectives for Medical Student Education. Guidelines for Medical Schools*: Association of American Medical Colleges; 1998.

38. ACGME. The Outcomes Project. http://www.acgme.org/outcome/. Accessed 20 June, 2006.

39. ACGME. Toolbox of Assessment Methods. http://www.acgme.org/outcome/assess/toolbox.asp. Accessed 20 June, 2006.

40. Greiner AC KE. *Health Professions Education: A Bridge to Quality*. Washington D.C.: Institute of Medicine; 2003.

41. ABIM Foundation. American Board of Internal Medicine A-AFACoP-ASoIM, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter.[see comment]. *Annals of Internal Medicine*. Feb 5 2002;136(3):243-246.

42. Blank L KH, McDonald W, Merino J, ABIM Foundation, ACP Foundation, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter 15 months later. *Annals of Internal Medicine*. May 20 2003;138(10):839-841.

43. Stern DT FA, Gruppen LD. The prediction of professional behaviour. *Medical Education*. Jan 2005;39(1):75-82.

44. Ginsburg S RG, Stern D, Lingard L. The anatomy of the professional lapse: bridging the gap between traditional frameworks and students' perceptions. *Academic Medicine*. Jun 2002;77(6):516-522.

45. Arnold L. Assessing professional behavior: yesterday, today, and tomorrow. *Academic Medicine*. Jun 2002;77(6):502-515.

46. Ginsburg S RG, Hatala R, McNaughton N, Frohna A, Hodges B, Lingard L, Stern D. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Academic Medicine*. Oct 2000;75(10 Suppl):S6-S11.

47. Papadakis MA OE, Cooke M, Healy K. A strategy for the detection and evaluation of unprofessional behavior in medical students. University of California, San Francisco School of Medicine Clinical Clerkships Operation Committee. *Academic Medicine*. Sep 1999;74(9):980-990.

48. Epstein RM HE. Defining and assessing professional competence.[see comment]. *JAMA*. Jan 9 2002;287(2):226-235.

49. Stern DT. *Measuring Medical Professionalism*. New York: Oxford University Press, Inc.; 2006.

50. Garfinkel PE DB, Sadavoy J, Bagby RM. Boundary violations and departments of psychiatry. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie.* Sep 1997;42(7):764-770.
51. ACGME. Table of Toolbox Methods. September 2000. Accessed 20 June, 2006.
52. Tamblyn RM KD, Schnable GK, Kopelow ML. Sources of Unrealiability and Bioas in Standardized-Patient Rating. *Teaching and Learning in Medicine.* 1991;3(2):74-85.
53. Pangaro LN W-DH, Macmillan MK, Klass DJ, Shatzer JH. Performance of "standardized examinees" in a standardized-patient examination of clinical skills. *Academic Medicine.* Nov 1997;72(11):1008-1011.
54. Tamblyn RM KD, Schnabl GK, Kopelow ML. The accuracy of standardized patient presentation. *Medical Education.* Mar 1991;25(2):100-109.
55. Sanson-Fisher RW PA. Simulated patients and the assessment of medical students' interpersonal skills. *Medical Education.* Jul 1980;14(4):249-253.
56. Norman GR TP, Feightner JW. A comparison of resident performance on real and simulated patients. *Journal of Medical Education.* Sep 1982;57(9):708-715.
57. Hodges B RG, Hanson M, McNaughton N. An objective structured clinical examination for evaluating psychiatric clinical clerks. *Academic Medicine.* Aug 1997;72(8):715-721.
58. Arnold L SD. Content and Context of Peer Assessment. In: DT S, ed. *Measuring Medical Professionalism.* First ed. New York: Oxford University Press, Inc.; 2006:175-194.
59. Shue CK AL, Stern DT. Maximizing participation in peer assessment of professionalism: the students speak. *Academic Medicine.* Oct 2005;80(10 Suppl):S1-5.
60. Woolliscroft JO HJ, Patel BP, Swanson DB. Resident-patient interactions: the humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors, and nurses. *Academic Medicine.* Mar 1994;69(3):216-224.
61. Burack JH ID, Carline JD, Root RK, Larson EB. Teaching compassion and respect. Attending physicians' responses to problematic behaviors. *Journal of General Internal Medicine.* Jan 1999;14(1):49-55.
62. Irby D MS. The legal context for evaluating and dismissing medical students and residents. *Academic Medicine.* Nov 1989;64(11):639-643.
63. Board of Curators of the University of Missouri v. Horowitz, 98 S.Ct. 948, 435 U.S. Supreme Court 1978
64. Dixon vs. Alabama Board of Education, 294 F.2d 150, 5th Circuit Court 1961
65. UT Southwestern Medical School Catalog. http://www8.utsouthwestern.edu/utsw/cda/dept24675/files/41199.html. Accessed 23 June 2006, 2006.
66. University of Texas System Board of Regents' Rules and Regulations. http://www.utsystem.edu/bor/rules/CompleteTOC-2.htm#studentissues. Accessed 23 June, 2006.
67. Connelly v. University of Vermont, 562 F.Supp. 156, Vermont District Court 1965

68. National Labor Relations Board v. Knoxville Publishing Co., 124 F.2d 875, 6th Circuit Court 1942
69. McDonald v. Hogness, 92 Wn.2d 431, Washington Supreme Court 1979
70. Regents of the University of Michigan v. Ewing, 106 S.Ct. 507, U.S. Supreme Court 1985
71. Veloski JJ FS, Boex JR, Blank LL. Measuring professionalism: a review of studies with instruments reported in the literature between 1982 and 2002. *Academic Medicine.* Apr 2005;80(4):366-370.
72. Ackoff RL. *Redesigning the Future - a Systems Approach to Societal Problems*: John Wiley & Sons, Inc. New York; 1974.
73. Ackoff RL. Our Changing View of the World. In: Ackoff RL, ed. *Ackoff's Best: His Classic Writings on Management.* New York: John Wiley & Sons, Inc.; 1999.
74. Veloski J HM. Measuring Specific Elements of Professionalism: Empathy, Teamwork, and Lifelong Learning. In: DT S, ed. *Measuring Medical Professionalism.* New York: Oxford University Press, Inc.; 2006.
75. Hochschild AR. *The Managed Heart: Commecialization of Human Feeling.* Second ed. Berkeley and Los Angeles: Univesity of California Press, Ltd.; 2003.
76. Larson EB YX. Clinical empathy as emotional labor in the patient-physician relationship.[see comment]. *JAMA.* Mar 2 2005;293(9):1100-1106.
77. Finestone HM CD. Acting in medical practice.[see comment]. *Lancet.* Sep 17 1994;344(8925):801-802.
78. Knaus WA WD, Zimmerman JE, Draper EA. Variations in mortality and length of stay in intensive care units.[see comment]. *Annals of Internal Medicine.* May 15 1993;118(10):753-761.
79. Hojat M. Philadelphia; 2006:Personal communication to Wagner JM.
80. Physician Achievement Review. http://www.par-program.org/PAR-Info.htm. Accessed 23 June, 2006.
81. Lipner RS Bylsma WH AG, Fortna GS, Tooker J, Cassel CK. Who is maintaining certification in internal medicine--and why? A national survey 10 years after initial certification.[see comment][summary for patients in Ann Intern Med. 2006 Jan 3;144(1):I36; PMID: 16389248]. *Annals of Internal Medicine.* Jan 3 2006;144(1):29-36.
82. Lipner RS BL, Leas BF, Fortna GS. The value of patient and peer ratings in recertification. *Academic Medicine.* Oct 2002;77(10 Suppl):S64-66.
83. Ramsey PG CJ, Inui TS, Larson EB, LoGerfo JP, Wenrich MD. Predictive validity of certification by the American Board of Internal Medicine.[see comment]. *Annals of Internal Medicine.* May 1 1989;110(9):719-726.
84. Sharp LK BP, Lipsky MS, Horowitz SD, Miller SH. Specialty board certification and clinical outcomes: the missing link. *Academic Medicine.* Jun 2002;77(6):534-542.
85. Cassel CK HE. Credentialing and public accountability: a central role for board certification.[comment]. *JAMA.* Feb 22 2006;295(8):939-940.
86. Freed GL SD, Lakhani I, Wheeler JR, Stockman JA 3rd, Research Advisory Committee of the American Board of Pediatrics. Use of board certification and recertification of pediatricians in health plan credentialing policies.[see comment]. *JAMA.* Feb 22 2006;295(8):913-918.

87. Freed GL UR, Hudson EJ, Lakhani I, Wheeler JR, Stockman JA 3rd, Research Advisory Committee of the American Board of Pediatrics. Policies and practices related to the role of board certification and recertification of pediatricians in hospital privileging.[see comment]. *JAMA*. Feb 22 2006;295(8):905-912.

88. Schneider ME. Grandfathered Physicians: A Few Choose to Recertify. *Internal Medicine News*. December 15, 2005, 1995: 60.

89. Baron RJ. Personal metrics for practice--how'm I doing? *New England Journal of Medicine*. Nov 10 2005;353(19):1992-1993.

90. Brennan TA HR, Duffy FD, Cassel CK, Goode LD, Lipner RS. Recertification for internists--one "grandfather's" experience. *New England Journal of Medicine*. Nov 10 2005;353(19):1989-1992.

91. Testerman JK MK, Loo LK, Worthley JS, Lamberton HH. The natural history of cynicism in physicians. *Academic Medicine*. Oct 1996;71(10 Suppl):S43-45.