Assessment of Professionalism : Recommendations from the Ottawa 2010 Conference

BRIAN DAVID HODGES[1], SHIPHRA GINSBURG[1], RICHARD CRUESS[2], SYLVIA CRUESS[2], RHENA DELPORT[3], FRED HAFFERTY[4], MING-JUNG HO[5], ERIC HOLMBOE[6], MATTHEW HOLTMAN[7], SADAYOSHI OHBU[8], CHARLOTTE REES[9], OLLE TEN CATE[10], YUSUKE TSUGAWA[11], WALTHER VAN MOOK[12], VAL WASS[13], TIM WILKINSON[14] & WINNIE WADE[15]

[1]University of Toronto, Canada, [2]McGill University, Canada, [3]University of Pretoria, South Africa, [4]Mayo Clinic, USA, [5]National Taiwan University, Taiwan, [6]American Board of Internal Medicine, USA, [7]National Board of Medical Examiners, USA, [8]Rikkyo University, Japan, [9]University of Dundee, UK, [10]University Medical Center, Netherlands, [11]St Luke's International Hospital, Japan, [12]Faculteit Health, Medicine and Life Sciences, Netherlands, [13]Keele University School of Medicine, UK, [14]University of Otago, New Zealand, [15]Royal College of Physicians, UK

Abstract

Over the past twenty-five years, professionalism has emerged as a substantive and sustained theme, the operationalization and measurement of which, has become a major concern for those involved in medical education. However, how to go about establishing the elements that constitute *appropriate* professionalism in order to assess them is difficult. Using a discourse analysis approach, the International Ottawa Conference Working Group on Professionalism (IOCPWG) studied some of the dominant notions of professionalism, and in particular the implications for its assessment. The results presented here reveal different ways of thinking about professionalism that can lead toward a multi-dimensional, multi-paradigmatic approach to assessing professionalism at different levels: individual, inter-personal, societal-institutional. Recommendations for research about professionalism assessment are also presented.

Background

The Theme of Professionalism

Over the past twenty-five years, professionalism has emerged as a substantive and sustained theme within both clinical medicine and medical education. Featured in medical education conferences and journals, the definition, operationalization and measurement of professionalism has become a major concern for those involved in the education and development of medical students as well as residents (house officers, foundation year doctors, etc), fellows, faculty, clinicians and researchers. And yet it is a topic with much ambiguity, confusion and at times controversy. The idea that the medical profession should attend to the professional behaviour of students and practitioners is not in dispute. However, how to go about establishing the elements that constitute

appropriate professionalism is more difficult. Though myriad studies have addressed this topic, the question: "what is professionalism?" remains complex and defining best practices for its assessment even more so. Difficulty stems from the notion that professionalism stretches along a continuum from the individual (attributes, capacities and behaviours) through the interpersonal domain (interactions with other individuals and with contexts) to the macro-societal level where notions such as social responsibility and morality but also political agendas and economic imperatives reside. Furthermore, there are interactions amongst these domains. For example an individual's professional behaviour may be influenced by the context; similarly, the individuals within an institution may influence its collective professional values.

While discussions and research about professionalism have appeared most prominently in Anglo-Saxon medical education literature in the past two decades, the globalization of medical education means increasing interest in the construct of professionalism in other languages, countries and cultures. As professionalism is a complex and multi-dimensional construct, one should not look for one, simple, generalizable statement about what professionalism is and how to assess it. Rather, assessment of professionalism requires consideration of its individual, inter-personal and societal dimensions.

Method

The International Working Group

The International Ottawa Conference Working Group on the Assessment of Professionalism (IOC-PWG) was created. Individuals with a history of writing or speaking about the topic internationally were invited to participate. Attention was given to diversity. The final group consisted of 18 individuals representing 9 countries in North America, Europe, Asia, Oceania and Africa. Seven members were women, 11 were men and represented a mix of clinicians and PhD-level academics from different disciplines (sociology, psychology, education, medicine, etc).

Discourses of Professionalism: Implications for Assessment

The working group chose to define the key issues related to the assessment of professionalism by undertaking a discourse analysis. We began with a set of articles identified by the IOC-PWG as key to the consideration of the assessment of professionalism. A discourse analysis is quite different from a traditional review in that the goal is not simply to summarize and condense existing findings, as would be done in a meta-analysis or summary review paper, but instead to characterize different ways that language is used to talk about and create statements of truth about a given phenomenon. There are many approaches to discourse analysis. The approach used here is inspired by what is known as critical discourse analysis.[1] In our discourse analysis of literature on professionalism, it was not the purpose to identify all papers on the topic, or to try to reduce findings down to a single set of consensus statements. Rather, the objective was to identify several discourses that are currently used to frame what professionalism is, and to form guidelines for how professionalism might be assessed. While a discourse analyst

tries to identify and classify samples of writing/text into a limited number of conceptual categories, it is important not to reduce or synthesize them to the point that paradigmatic nuances are blurred.

A discourse analysis is particularly well suited for something as complex and multi-faceted as professionalism. Categorization helps to illustrate the diversity of active discourses related to professionalism.[1] There is no assumption that these are the *only* categories nor that they would be fixed over time or in different places. The purpose of this discourse analysis is to reveal different ways of thinking about professionalism so as to allow researchers, educators and clinicians to preserve their core values, interests and paradigmatic perspectives and at the same time collectively to work toward a multi-dimensional, multi-paradigmatic approach to assessing professionalism.

Discourse Analysis of Key Articles

The 18 members of the international working group each submitted 2-3 references (original research, theoretical article, review paper, etc) that they considered "key articles in assessment of professionalism". A few redundancies were eliminated and 50 articles were downloaded. Articles were then read in detail by the group lead (BH). Papers were coded for key words, concepts, and an anchoring/representative statement about the nature of professionalism for each of the articles was identified. Specific implications for assessment from each of the articles were extracted. Articles were sorted into groups according to similar discourses/statements about the nature of professionalism and its assessment. The preliminary classification was shared with working group members who provided feedback through an iterative approach of subsequent drafts refined and re-circulated repeatedly over several months.

Implications for assessment were summarized from papers dealing with each level of discourse about professionalism. As well, potential limitations/weaknesses/implications of thinking about professionalism using each discourse were considered. An anchoring concept of the work was that no one discourse would encompass every dimension of professionalism and that there was benefit in understanding what might be left out or obscured through using only one of the discourses. In the words of Kenneth Burke, "Every way of seeing is also a way of not seeing". Draft recommendations were created through an iterative process involving all members of the working group.

The draft recommendations were presented in multiple venues at the International Ottawa Conference in Miami in May 2010, the Association for Medical Education in Europe conference in Glasgow in August 2010 and Association of American Medical Colleges in Washington in November 2010. They were also posted on the website of the International Ottawa conference for comment. All sources of feedback were used to make final refinements to the recommendations.

[1] The terms *discourse, epistemology,* and other terms used in this analysis are defined in a glossary shown in Appendix 1.

Results

Classification of professionalism discourses by scope (individual, interpersonal, societal/institutional) and epistemology (objectivist/positivist or subjective/constructivist) (See Table 1 for definitions)

Articles about professionalism were classified according to the different discourses used by authors, underpinning their perspective on what professionalism is, how its nature can be discovered and whether or not they believe it to be relatively constant across time and cultures or something that is highly changeable. Table 1 is an orientation matrix to the way in which the various professionalism discourses were grouped together. The organization follows two dimensions: scope (individual, interpersonal, societal/institutional) and epistemology (objectivist/positivist or subjective/constructivist). These terms are explained in the glossary in Appendix 1. It is important to note that these are not fixed, discrete categories. Rather they should be considered to represent continua. The levels of scope – individual, interpersonal and societal/institutional - overlap and also represent a continuum from the individual to the collective. The epistemological "positions" described in this table can be thought of as *dominant* perspectives or "leanings" toward a certain view of how the world works. There were, in some instances, tensions or contradictions between positions and authors of papers (and members of the working group) often moved between perspectives.

Insert Table 1 about here

Having read and classified all of the key articles, and drawing on the collective expertise of the 18 members of the international working group, three overarching discourses about professionalism assessment were identified: individual, interpersonal and societal/institutional. From these the following General Principles relating to the assessment of professionalism were developed iteratively over a 6-month period:

1. Professionalism is a concept that varies across historical time periods and across cultural contexts.

2. The need to develop concrete and operationalizable definitions, and from them effective teaching methods and defensible assessment approaches across the continuum of professional development, is strongly felt by many medical educators.

3. Professionalism is intrinsically related to the social responsibility of the medical profession. Thus, developing an acceptable, clearly articulated and operationalizable definition that is reviewed and refined regularly to reflect societal and health care changes, is an important responsibility of the profession and its educational institutions to the public.

4. What professionalism is and how it will be taught and assessed should be clearly articulated through a dialogue between the profession and the public. Professionalism

can be conceptualized and assessed at different levels: individual, interpersonal and institutional/societal. A comprehensive understanding of professionalism requires attention to these multiple, and often interdependent, levels.

5. A culture that fosters continual improvement of all students and practitioners, and emphasizes personal and collective responsibility for that improvement is desirable. While summative assessment is important, formative methods should predominate including robust feedback for all students and practitioners, supplemented where necessary by remediation.

6. Professionalism, and the literature supporting it to date, has arisen predominantly from Anglo-Saxon countries. Caution should be used when transferring ideas to other contexts and cultures. Where assessment tools are to be used in new contexts, re-validation with attention to cultural relevance is imperative.

7. Different perspectives lead to different statements about the nature of professionalism. They represent different lenses and focus attention on different aspects of education, assessment and research in this domain. A diversity of approaches and perspectives (psychometrics, psychology, sociology, anthropology, etc) should be embraced in professionalism assessment and research.

8. Each perspective (and resulting assessment methods) will make some elements of professionalism visible, and each will deflect attention from other elements. Elements of professionalism are vast and include: individual (attributes, characteristics, attitudes, behaviours, identities), interpersonal (relations, group dynamics, etc) and societal (economic, political, etc).

Having defined some general recommendations about professionalism the group then turned to defining key issue for assessment at each of the three discourses.

Three discourses about professionalism and recommendations for assessment.

1. Professionalism as an individual characteristic, trait, behaviour or cognitive process

In this discourse, professionalism is understood to exist or develop to varying degrees as a *characteristic* or *attribute* that is identifiable within individuals. Working within this discourse means focusing on the individual: attending to, and prioritizing, their attributes, whether believed to be inherent (essentialist) or mutable (developmental/learned). Significant attention is given to the measurement of these attributes, usually in the psychometric tradition. The context in which the attributes are expressed is less of a focus, and there is generally an assumption that the attributes are relatively stable and can be captured by tools that are sufficiently valid and reliable. The distinction between an essentialist perspective and a developmental perspective is not sharp, with some authors allowing for the presence of both elements. In addition, some attributes are considered to be more stable (traits) than others (states).

Authors working with an essentialist perspective view professionalism as a set of inherent personality traits apparent prior to admission to medical school (and therefore relatively fixed). They argue that diagnostic screening tools are necessary at the time of selection for admission to medical school.[2] They suggest that standardized instruments are needed to assess the personal qualities of medical school applicants that predict problematic performance; also need is an improved system of evaluation to document deficiencies and that provides remediation, is central.[3] However, one study reported that there were no consistent, significant correlations between any materials from the admissions packet and any of the outcomes of professional behaviour by year 3 of medical school, although missing immunizations, missing evaluations and self-assessment appeared to correlate with professionalism ratings.[4] Another paper suggested there was a relationship between professionalism as estimated by medical students' peers and an index of "conscientiousness".[5]

Principles distilled from such papers are that some component of professionalism may be related to inherent personality characteristics or traits. Assessment of traits (cognitive, personality, behavioural, etc) prior to admissions may be relevant to later professionalism, but this remains speculative. Links still need to be shown between pre-admissions data, medical school performance, residency performance and professionalism in practice. Cautions raised by authors working within this paradigm about this approach are that research has not yet identified specific characteristics or traits that robustly predict future behaviours from the premedical period. However, more evidence is available about the link between medical school performance and behaviour in practice. Concerns associated with false positives/negatives are raised in relation to high stakes measurement, as well as hesitations about "homogenization" given the desire for a diverse student population that will serve different roles/purposes in practice.

The Working Group identified a tension between those who wish to teach professionalism as essentially a moral endeavor and those who wish to have a list of attributes. Although many papers tend towards the list of traits approach, there are individuals, often writing from a background in ethics, psychology or sociology of the professions who are uncomfortable with a trait theory, or a "personality" approach.

Also using the individual discourse, but taking a somewhat more developmental/educational approach is a set of papers that focus on professionalism as learned behaviours that develop *during* medical education. Several of these focused on the use of the "Professional Mini-Evaluation Exercise (PMEX)" a four-factor, 24-item instrument with sufficient validity/reliability with approximately 8 raters.[7] Another measure of observable behavior reviewed was the Amsterdam Attitude and Communication Scale (AACS).[8] A third set of papers focused on Deans' letters and their content about professional behaviour.[9] Together, these papers argued for the need to clarify elements of professionalism and to develop better tools to assess behaviours (psychometrics) by peers, teachers and during critical incidents. One author suggested improving assessment by anchoring the assessed behaviours in real-world value conflicts.[10] Some papers underlined the need to create systems to foster peer feedback by emphasizing anonymity, immediacy, ubiquity, documentation, formative approaches

(punishment/correction, "hold students responsible") for unprofessional behaviours and to "reward" professional behaviours.[11] A challenge to these approaches is that measures of observed behaviours, self-reports and single attributes are not considered adequate to assess professionalism by some authors who argue for the need to develop measures of values and attitudes and understand their relation to behaviour change.[12] Finally it was argued that there are many existing assessment scales and ratings (one author reported finding 88 of them)[13] and that existing measures should be improved psychometrically, rather than continually creating new ones. Others called for including many assessors, more than one assessment method and assessment in different settings such as multi-source feedback, cognitive assessments and patient questionnaires. [13]

Some authors taking an individualist approach focus on the postgraduate level. For example it was shown that residency professionalism ratings and written exams (ABIM certification exams) can predict some future problem behaviours.[14] At this level some have argued that most tools are designed to evaluate specific elements of professionalism but that few assess a comprehensive construct. One paper recommended that multi-source feedback should not be the sole measure of professional behaviours: "A pragmatic approach is needed whereby multiple snapshots of an individual's professionalism can be taken and collated into a whole to develop a clear picture of that person's strengths and weaknesses and to provide a body of evidence on which to base summative decisions".[15]A complex, multi-tool blueprint therefore is required. One study found that formal evaluation sessions (verbal discussions) actually contained more references to unprofessional behaviours than checklists or rating forms.[16] A final piece in relation to professionalism as an individual characteristic is that both student wellbeing and professional behaviours should be monitored continuously and rigorously with a system of data collection, analysis, interpretation and intervention and that it is important to be clear whether the system is supportive or regulatory or one that combines elements of both.

Overarching principles distilled from the papers in this group include the notion that professionalism may be understood as the observable, behavioural manifestations of the interaction of a complex set of cognitive, attitudinal, personality and characteristics. This makes clear that the assessment of behaviours is a proxy measure, resting on the assumption that these behaviours are fully (or at least significantly) reflective of the underlying dimensions of professionalism. Thus, in order to be fair and defensible, the assessment of behaviours should be done using instruments that have demonstrable reliability and validity. Documenting behaviours alone, however, may be insufficient to capture a comprehensive construct of professionalism that also includes knowledge, attitudes and the ability to employ professional behaviours in real practice settings. Some have argued that by focusing on behaviours it is frequently forgotten that one can test a student's knowledge of professionalism. Professionalism has a knowledge base and including it in the subject matter to be tested will drive learning as it does in other areas. This is rarely reflected in the literature on assessing professionalism, which concentrates on behaviors. Overall, the best assessments are part of a program that includes setting a safe climate, feedback, anonymity when appropriate and follow-up of behaviour change as documented by several measurements over multiple time periods. Finally, it appears

that identification and documentation of "negative" behaviours may be distinct, and in the minds of some less important than systems that recognize and document "positive" professionalism behaviours.

Cautions raised about this individual discourse focus mainly on the idea that the assessment of behaviours alone may not reflect contextual aspects of professionalism. Observable behaviours may have more to do with the exigencies of particular contexts than of deeply held values and attitudes. In other words, behaviours may be highly unstable across different contexts. There are aspects of professionalism that may be obscured by focusing on the individual. Students and teachers often struggle to define what professionalism means to them and note that what they consider "professional" in one setting may not be in another. By downplaying the importance of context, perfectly reasonable students can sometimes be demonized as "unprofessional" rather than just having "lapsed" due to time pressures, hierarchical pressures, etc. Further, if tools are created for specific contexts (institutions, specialties, cultures, countries) students and teachers may not value definitions or constructs of professionalism that feel "imported". For example, those writing about professionalism in Asian countries have noted a "buy-in" problem when definitions of professionalism and assessment tools are simply translated from North American versions. There are also generational issues that relate to the interpretation of behaviours vis-à-vis such concepts as "altruism" and "lifestyle". Trying to teach what Hafferty calls "nostalgic professionalism"[17] may result in simple rejection by the current generation.

The following recommendations were elaborated for assessment of professionalism as an individual phenomenon.

1. Some component of professionalism may be related to inherent personality characteristics or traits. Assessment of traits (cognitive, personality, behavioural) prior to admissions may be relevant to later professionalism; however, use of such screening approaches requires that links between pre-admissions data, medical school performance, residency performance and professionalism-in-practice be demonstrated.

2. Professionalism may be understood as the external, behavioural manifestations of the interaction of a complex set of cognitive and attitudinal elements and personality characteristics, mutually and with the environment. However behavioural assessments are proxy measures, resting on the assumption that observed behaviours are reflective of underlying dimensions. Research shows that this assumption is not always accurate. For this reason, documenting behaviours alone may be insufficient to capture a comprehensive construct of professionalism, which should also include knowledge, values, attitudes and the ability to employ professional behaviours in real practice settings.

3. Where behavioural assessments are used, instruments should be employed that have demonstrable reliability and can be used to support valid inferences. Both quantitative measures (e.g. numeric scores derived from observation-based survey instruments)

and qualitative measures (e.g. narrative data from Dean's letters) have been studied and may be employed in a defensible manner. A combination of methods over a period of time is likely to be needed.

4. Given the number of existing professionalism assessment tools, it may be more important to increase the depth and quality of the reliability and validity of a program existing measures in various contexts than to continue developing new measures for single contexts.

5. Triangulation of multiple kinds of measures, by multiple observers, synthesized over time with data gathered in multiple, complex and challenging contexts is likely to be appropriate at all levels of analysis.

6. Identification and documentation of *negative* behaviours is likely to require a distinct system from one in which there is recognition, documentation and reinforcement of *positive* professionalism behaviours. Instrument design and validity research should be undertaken thoughtfully in such a way as to reflect this distinction.

7. The overall assessment program is more important than the individual tools. The best programs use a variety of tools in a safe climate, provide rich feedback, anonymity (when appropriate) and follow-up of behaviour change over time. Effective assessment and feedback programs also incorporate faculty development.

2. Professionalism as an interpersonal process or effect

In this discourse professionalism is understood to be something constructed (or suppressed) through inter-personal interaction. Working in this discourse means giving attention to interpersonal relationships, particularly that of student and teacher. While individual attributes are still a focus, these are understood to be co-created between a student and another person (teacher, patient, etc) and therefore more fluid. Context is given significant attention, as is the notion that the expression of professionalism is contextually determined. The detection and assessment of professional behaviours cannot take place without an analysis of the context in which they are expressed. Writers working with this discourse often express greater interest in formative assessment for teaching and learning, and somewhat less focus on summative assessment, but this need not be the case. The context, student-teacher, student-student and student-health professional relationships and the learning climate itself may be targets for assessment as much, or more so, than individuals.

What this discourse makes visible/possible is the identification, documentation and analysis of relationships on student and teacher perceptions of professionalism, and attention to context. On the other hand, this discourse can obscure macro-social forces acting on the teacher-student dyad and the institution in which learning occurs. It also gives less attention to personality attributes/traits and may not be as helpful in finding ways to address the rare but problematic individuals. Overly focusing on contextual dimensions might also diminish a sense of personal responsibility among students.

There are many variations on the interpersonal approach to professionalism. For example, studies have examined the idea that professionalism is a set of socio-cognitive processes that an individual uses to interpret problems in the world and to select responses in relation to others. Ginsburg and colleagues for example, set out to find generalizable cognitive features of problem solving that might shed light on the cognitive processes behind observed behaviours. They argued that, as no fixed list of traits could be defined, nor could raters be standardized, assessment should involve exposing students to value conflicts and having them produce a resolution, observing and scoring the process they use, the values and principles invoked and the decisions made. They introduced the concept of "professionalism lapse" as more useful than the label "unprofessional".[18] They wrote that: "Future efforts at evaluation need to look beyond the behaviors, and should incorporate the reasoning and motivations behind students' actions in challenging professional situations…sophisticated evaluation of professionalism requires an additional dimension, as behaviors alone do not give us all of the information we need to make accurate judgments."[19,20] Others have argued that there are definable stages that individuals pass through on the way from "proto" professionalism to full professionalism in relation to learning environments. Evaluation involves the documentation of attainment (or attrition) of these characteristics.[21] To do this, reliable and valid ways to characterize the learning environment are needed. According to these authors, institutions should measure and maintain high professional standards of the learning environment. Initiatives to improve professionalism should be evaluated in terms of their impact on the environment.[22]

Taken together, principles distilled from these papers are that there are common features of unprofessional behaviour/professionalism lapses that arise from particular kinds of social interactions and that these are generalizable across contexts. Assessment should include exploration of students' cognitive problem solving processes, monitoring learning environments as well as teacher-student relationships for interpersonal characteristics that could lead to unprofessional behaviours/professionalism lapses. Cautions voiced by authors working within this discourse include the idea that broadening the perspective to include teachers and the environment can be threatening to teachers. What using this discourse may obscure is that the nature of these inter-personal effects may be specific to cultures (by country, ethnicity, tradition or even institutional).

A somewhat more constructivist approach begins with the notion that professionalism is a way of being that is entirely created in interpersonal interactions. According to this perspective, behaviour results from the generation and negotiation of meaning through interaction with others. This view draws on social psychology, symbolic interactionism and developmental psychology. For example, it is argued that professionalism is subtle and complex and does not reduce to numerical scales; that most assessment overemphasizes factors related to the person and underemphasizes factors related to the context. Some recommend exploring assessment that does not rely on scales at all.[20] As one author put it, the implication is that measurement of the student alone is only half of the equation.[23] The key point is that relying on behavioural assessment might lead to passing students with "professional behaviours" but unethical attitudes and fail students

with "unprofessional behaviours" but ethical attitudes. Thus assessment must include context-dependent nature of behaviours. Observation alone is not enough. Conversations about behavior, and behavioural explanations, are key. Thus it is necessary to collect data using multiple methods including observations and interviews and focus on text and narrative.[24] A central idea here is that assessors have a role in constructing students' unprofessional behaviours.[25] The environment should therefore be monitored for conditions that lead to negative phenomena such as the emotional detachment of students.[23]

A key principle distilled from these papers is that professionalism is a set of behaviours and responses to situational and contextual phenomena that arise much less from individual cognitive or personality dimensions and much more from context during learning and practice. The assessment of professionalism therefore involves assessing the thoughts, decisions, responses and behaviours of all actors in each context, perhaps most importantly both teacher and student. Assessment of the learning/practice environment itself is also important. Inherent in this approach to assessment is feedback to improve the performance of individuals (teachers, other health professionals) and of the context/learning environment itself. The concept of "unprofessionalism" (a characteristic or trait) is less useful than "professionalism lapses" (situation).

Cautions raised by authors working within this discourse include that assessing characteristics and behaviours of students alone, without an assessment of other members of the system and of the context itself risks missing important forces that shape and determine behaviour. It is important to make the connection between a necessarily reductionist set of observable behaviors and something more profound, and necessarily subjective. What may be obscured by a focus on this discourse of professionalism is how difficult it is to conceive of any program of evaluation of student's knowledge of professionalism and of professional behaviors that does not start with something fairly concrete. From this perspective, the need to define universal features of professionalism (e.g. "*primum non nocere*" or "patient interest above personal interest") may be strongly felt.

The following recommendations were elaborated for assessment of professionalism as an interpersonal phenomenon.

1. In addition to its individual elements, professionalism also implies a set of behaviours and responses to situational and contextual phenomena that arise during learning and practice. The assessment of professionalism should therefore include assessment of the decisions, responses and behaviours of all actors in each context (perhaps using multi-source feedback), gathering longitudinal data from both teacher and student as well as other key players such as health care professionals, administrators, patients, etc.

2. Assessment of the learning/practice environment itself is also important. Inherent in this approach to assessment is feedback to improve the performance of teams (course faculty, clinical teaching teams, etc.) as well as to improve structural elements, be

they organizational (e.g., policies that govern learning/work) or structural in an architectural sense.

3. Assessment of professionalism should include monitoring learning environments, student-student, teacher-student, student-health professional and student-patient relationships for problematic interpersonal phenomena. The concept of situationally specific *professionalism challenges, dilemmas* or *lapses* may be more useful than a global concept of *unprofessionalism* (characteristic or trait).

4. While complete consensus on what are appropriate professional responses to complex problems and situations may not always be achieved completely, assessment and feedback should represent a collective perspective where possible.

3. Professionalism as a societal/institutional phenomenon: A socially constructed way of acting or being, associated with power

In this discourse, a key notion is that professionalism emerges and is modified through the interaction of professional groups with society. Professionalism is something that serves a social purpose of some higher order. That is, professionalism has a function – be it in relation to the status of the profession, the organization of the health care system, or the cultural, social or moral structure of institutions and societies of which medicine is a part. In this sense, professionalism is defined with and by society. Individual attributes and inter-personal processes are inseparable from consideration of these larger forces but the emphasis is at the macro level.

There were two variations on this discourse in the papers reviewed. The first, an objective/positivist historical or utilitarian orientation, starts from the assumption that an objective professionalism exists and is relatively independent of context, generalizable and therefore shaped by, but not wholly created by, social forces. Assessment means tying together attributes and behaviours of individuals, but also of teams and professional groups, to outcomes at organizational, systems or social levels. Assessment is more likely to take the form of macro/social or institutional outcomes (patient outcomes) or processes (accreditation). What this discourse makes visible/possible is identification, documentation and analysis of socio-organizational elements and functions of professionalism for evaluation of efficiency, productivity, relevance or quality of medical professional practice and organization, and patient safety. What this discourse can obscure is the dynamics of power that construct particular definitions of what professionalism is in different times and places.

For example it is argued that professionalism is an aspect of identity, status and autonomy of the medical profession, drawing on systems theory and the study of professions. An implication is that medical schools, medical educators and the profession in general must emphasize setting expectations, teaching and assessing professionalism at a high level across the profession as a whole.[26] A related notion is that professionalism is a collective responsibility of the medical profession that arises from its social contract

with society with the implication that measurement should include the key elements outlined in the model. Both macro-dimensions (the contributions of each partner to the social contract – medical profession, but also government, society, etc) and micro-dimensions (individual level comportment of physicians) need to be assessed. Cruess and Cruess for example, separate out the contextual/country-specific elements of the professional "social contract" and what they consider to be more universal dimensions of individual behaviour associated with "the healer".[27]

The related idea is that professionalism is a set of attitudes and behaviours linked to systems requirements of cost control, access to care, efficiency and quality (production imperative) of health care, notions that draw on politics, economics and business management literatures. The implication is that attitudes and beliefs expressed should be measured against actual behaviours, recognizing the often large gap.[24] An example given is the conflict of interest scenario in which a doctor who owns a private clinic faces professional dilemmas about continuity of care that may challenge espoused beliefs because of a particular healthcare context.[28] An interesting argument in this literature is that attention to and assessment of professional values is necessary to ensure medicine does not become a "trade".[29] Assessment of professionalism would thus focus more on what individuals do in relation to the system in which they work rather than an individual's autonomy or self-determination.

Others taking this macro-societal perspective have argued that professionalism is a collective core set of values and approaches tied to morality and anchored in specific philosophical/ethical/religious traditions. The implications include a "need to move beyond validity arguments that have been made for … traditional assessments targeted at cognitive competencies". There is a need for "thick description" to "interpret the flow of meaningful events from participants' perspectives", because "social reaction and conduct are inseparable".[30,24] Finally, adding a contextual element are papers that suggest that professionalism is a set of definable and measurable behaviours that vary across cultures. For example, whereas psychometric evaluation with the P-MEX was reliable and acceptable in the Japanese context, nevertheless new items were needed and different results were obtained in Japan than in a Canadian setting.[31] Similarly, Taiwanese researchers proposed an approach to construct a professionalism framework that accounts for historical and socio-cultural context. The framework they built shared similarities with western counterparts but differs in the centrality of self-integrity, harmonizing social roles, reflecting Confucius values.[32]

To summarize, principles distilled from papers using this macro-societal discourse are that professionalism is an aspect of, and must be understood in the context of, the goals, aspirations and exigencies placed upon the profession as a whole. Assessment involves characterizing those expectations and measuring the degree to which the profession (be it a subgroup such as students, a whole medical school, a professional practice group, or even the profession as a whole) meets those expectations. Assessment and research on assessment therefore may involve critiquing the dominance of certain ways in which those expectations are framed or enforced. Authors working with this discourse grapple to some extent with the profession as a whole and institutions as "actors" unto

themselves. They start from the premise that what happens at the macro level sets the stage for (and constrains) the ways in which individuals calibrate their own professional actions. Cautions raised by authors working within this paradigm are that the nature of the professionalism in the future will be strongly influenced by societal decisions relating to national health care systems and changes in self-regulation. What may be obscured by this discourse is that research has not yet established that the concept known as "professionalism" in the Anglo-Saxon countries/English literature exists or is fully understandable in other cultures and linguistic groups.

Also working at the macro-societal/institutional level, but taking more of a social constructivist/critical perspective, some authors start with the premise that there is no one fixed entity called professionalism in all places and historical periods. Rather it is a phenomenon created through discourse and power in certain places and times. For writers working from this perspective, the lack of cross-cultural validation of the concept raises concerns that perhaps professionalism as defined in the Anglo-Saxon literature might have a different nature, or possibly not even be understandable in a different language or culture. Working in this discourse means putting aside the notion that there are any fixed attributes or behaviours called professionalism that can be defined in the same way in all times and places. Rather, professionalism is something that has arisen in some places/cultures/time periods in concert with specific social forces/discourses/values. More focus is given to the processes that create different conceptions of professionalism (or make it possible to exist at all) than the actual attributes or behaviours of individuals or groups. Assessment, often qualitative, focuses on the meanings and attributions that individuals and groups give to their context and the ways in which their identity and certain of their behaviours are considered "professional" (or unprofessional) and how this determination is shaped by social forces/dynamics/power (culture, gender, socio-economic status, etc.). What this discourse makes visible/possible is the identification, documentation and analysis of dynamics of power that lead to particular constructions of professionalism. It also highlights both the productive and repressive effects of power, hierarchy and social organization and institutions. What this discourse can obscure is the sense of urgency felt by educators to classify positive/pro-social characteristics as well as problematic behaviours for the purposes of admission to medical school or pass/fail/remediation decisions during medical training.

The key argument in such work is that professionalism is a social construction. This approach draws on sociology, political economy, historiography and anthropology. Assessment of individual characteristics or behaviours is therefore seen as inadequate.[17] As a complex, adaptive system, assessment of professionalism should entail means of analyzing motivations and behaviours in context, at individual (the medical student/teacher), institutional (the medical school) and societal (the medical profession) levels.[33] Authors working with this approach argue that professionalism is too complex and nuanced to be captured by checklists of individual characteristics or behaviours alone. Social-contextual factors shape the expression of behaviours, which may or may not reflect attitudes and values of individuals, or even small groups (e.g. teacher-student).[24] They argue that strategies for screening for character traits during admissions processes are not robustly predictive and might not even be desirable given the need for

diversity. As a "distributed" phenomenon, professionalism should be assessed in terms of the function of groups, settings and institutions more than individuals.[34]

Principles distilled from this social-constructivist orientation include that assessment is a risky business because it is an act of power with the possibility to discriminate. Constructions of the definitions of what professionalism is are themselves subject to power relations, including the projects and agendas of social groups and institutions and may disguise problematic constructions. Assessment in this perspective is about gathering data to demonstrate equity and fairness in processes that discriminate between individuals and the accountability of professional groups and institutions as a whole. Cautions raised by authors working within this approach are that those accustomed to the objectivist/positivist orientation may find a social-constructivist perspective disorienting, and worry that constructivism means that all things are relative/of equal value. Those accustomed to a social-constructivist approach may find an objectivist-positivist orientation difficult, and worry about that effects of power are hidden behind apparent objectivity. What may be obscured by this approach is that the "earnest search" for a measurable and teachable phenomena articulated by front line teachers and evaluators seems difficult or impossible.

The following recommendations were developed for assessment of professionalism as institutional/societal phenomenon:

1. Professionalism can be understood in the context of the goals, aspirations and collective behaviours of healthcare and educational institutions and of the profession itself. Assessment involves characterizing societal expectations, through dialogue and meaningful input from public stakeholders, and measuring the degree to which the profession (be it a subgroup such as students, a whole medical school, a professional practice group, or even the profession as a whole) meets these expectations. Accreditation requirements at every educational level require teaching and evaluating of professionalism. Effectiveness should be measured in terms of clear institutional/societal outcomes.

2. Assessment may involve critiquing the dominance of certain ways in which expectations and practices are framed or enforced (cultural, generational, gendered, hierarchical, etc.) and should lead to improved institutional and organizational climate and practice.

5. Professional lapses may arise from particular kinds of social interactions and problematic organizational and institutional settings and politics. Examining and making explicit the *hidden curriculum* and tacit problematic organizational or institutional norms is important in assessing and contextualizing professional/unprofessional behaviours of students, teachers and institutions.

Implications for research on professionalism assessment

Finally, it was widely recognized in the papers reviewed by members of the International Working Group that further research on the assessment of professionalism is warranted. The following recommendations were elaborated in relation to research about professionalism assessment

1. Examine the concept of professionalism and its assessment across different linguistic and cultural contexts.

2. Compare the definitions and conceptions of professionalism assessment in medicine to those held by other healthcare professions.

3. Characterize which elements of professional behaviour are amenable to learning (and therefore remediation) and which may have a more immutable quality that are amenable to selection processes.

4. Examine links between the assessment of professionalism and other assessment initiatives such as quality of patient care.

5. Develop and evaluate means of incorporating patients' perspectives into the assessment of professionalism.

6. Explore professionalism assessment in complex clinical workplaces, including how individuals adapt to difficult or even dysfunctional systems and the gaps that arise between espoused values and actual practice.

7. Elaborate ways that assessment data can be used to change the culture of education and practice, in particular the *hidden curriculum*.

8. Consider what happens when expectations at an individual level conflict with those at the societal/organizational/institutional level, and what the resolution means for professionalism assessment.

9. Explore innovative ways to collect and analyze quantitative and qualitative methods of assessment data from mixed-methods approaches, paying particular attention to threats to validity inherent in different assessment methods.

10. Conduct outcome studies to examine the impact of curriculum (formal, informal and hidden) and other organizational interventions related to professionalism.

Conclusions

A common approach to developing consensus recommendations is to review a wide range of literature, consult with experts and work toward a shared set of guidelines or

"best practices". In tackling the domain of professionalism, it was obvious from the outset that no one unified consensus would be possible, nor desirable, given the diversity of ways in which the phenomenon is understood. Rather than trying to force the paradigmatic richness that characterizes professionalism research into on overly simplistic list of recommendations, the International Working Group on the Assessment of Professionalism chose a discourse analysis approach. This allowed us to unearth, categorize and represent three key discourses about professionalism – as an individual, inter-personal or societal-institutional phenomenon – discourses that are all in active use today. The strength of this approach is that we were able to create recommendations specific to each of the three main discourses identified. The obvious corollary is that no unified "statements of truth" about what professionalism is or how is should be assessed are made.

The working group found the use of discourse analysis challenging but ultimately gratifying because of the strength of this method is to retain and value diverse perspectives and at the same time emphasize that all approaches both illuminate and obscure what is "true" about professionalism. For those interested in the complex and important topic of professionalism, we hope that we have provided new insight as well as some helpful directions for both assessment and for future research.

Appendix 1: Glossary of Key Terms [35,1]

Constructivism: A belief about knowledge (epistemology) that asserts that the reality we perceive is constructed by our social, historical, and individual contexts, and so there can be no absolute shared truth.
Discourse: a set of statements/logical system of thought that attempts to articulate the essence of what professionalism is as employed in a given article or body of work.
Discourse Analysis: A methodology that analyses language to enable an understanding of its role in constructing the social world. Critical discourse analysis focuses on the macro level features of oral and written texts in their social contexts (as opposed to "linguistic discourse analysis," which includes the micro level analysis of grammatical features).
Epistemology: underlying conception of how knowledge comes to exist; a theoretical approach to knowledge.
Methodology: method of data collection/analysis linked to an epistemological perspective.
Objectivism: A belief about knowledge (epistemology) that asserts that there is an absolute truth or reality that can be discovered and that knowledge is objective and neutral.
Positivism: A theoretical framework that is guided by the search for the objective truth that will contribute to the progress of humankind.

Table 1: Classification of professionalism discourses by scope and epistemology

		Scope		
Epistemology		Individual	Interpersonal	Societal/institutional
Positivist-objectivist	Generalizable	Professionalism is an objectively definable phenomenon to be found in individuals, generalizable across cultural contexts	Professionalism is an objectively definable phenomenon to be found in interpersonal interactions, generalizable across cultural contexts	Professionalism is an objectively definable phenomenon to be found in social groups, generalizable across cultural contexts
	Limited generalizability	Professionalism is an objectively definable phenomenon to be found in individuals, but shaped by context	Professionalism is an objectively definable phenomenon to be found in interpersonal interactions, shaped by context	Professionalism is objectively definable phenomenon to be found in social groups, shaped by context
Subjectivist-constructivist orientation		Professionalism is subjectively constructed within individuals, arising from cultural context	Professionalism is an interpersonally constructed phenomenon, arising from cultural context	Professionalism is a socially constructed phenomenon, arising from cultural context

References

[1] Hodges B, Kuper A, Reeves S. Qualitative methodology: Discourse analysis. *BMJ*. 2008;**337:a879**:570-572.

[2] Janice A Knights, Kennedy BJ. Medical school selection: screening for dysfunctional tendencies *Medical Education*. 2006;**40**:1058-1064.

[3] Papadakis M, Teherani A, Banach M, et al. Disciplinary action by medical boards and prior behavior in medical school. *New England Journal of Medicine* 2005;**353**:2673-2682.

[4] Stern DT, Frohna AZ, Gruppen LD. The prediction of professional behavior. *Medical Education*. 2005;**39**:75-82.

[5] Finn G, Sawdon M, Clipsham L, McLachlan J. **Peer estimation of lack of professionalism correlates with low Conscientiousness Index scores.** *Medical Education*. 2009;**43**:960-967.

[6] Epstein RM. Assessment in medical education.[see comment]. *New England Journal of Medicine*. 2007;**356**:387-396.

[7] Cruess R, McIlroy J, Cruess S, Ginsburg S, Steinert Y. The professionalism mini-evaluation exercise: A preliminary investigation. . *Academic Medicine*. 2006;**81**::S74-78.

[8] ten Cate TJ, de Haes J. Summative assessment of medical students in the affective domain. *Medical Teacher*. 2000;**22**:40-43.

[9] Shea JA, O'Grady E, Wagner BR, Morris JB, Morrison G. Professionalism in Clerkships: An Analysis of MSPE Commentary. *Academic Medicine*. 2008;**83**:S1-S4.

[10] Arnold L. Assessing professional behavior: yesterday, today, and tomorrow. *Academic Medicine*. 2002;**77**:502-515.

[11] Arnold L, Shue CK, Kalishman S, et al. Can there be a single system for peer assessment of professionalism among medical students? A multi-institutional study. *Academic Medicine*. 2007;**82**:578-586.
[12] Jha V, Bekker HL, Duffy SR, Roberts TE. Perceptions of professionalism in medicine: a qualitative study. *Medical Education*. 2006;**40**:1027-1036.
[13] Lynch D, Surdyk P, Eiser A. Assessing professionalism: a review of the literature. *Medical Teacher*. 2004;**26**:366-373.
[14] Papadakis MA, Arnold GK, Blank LL, Holmboe ES, Lipner RS. Performance during Internal Medicine Residency Training and Subsequent Disciplinary Action by State Licensing Boards. *Annals of Internal Medicine*. 2008;**148**:869-876.
[15] Wilkinson TJ, Wade WB, Knock LD. A blueprint to assess profressionalism--results of a systematic review. *Academic Medicine*. 2009;**84**:551-558.
[16] Hemmer PA, Hawkins R, Jackson JL, Pangaro LN. Assessing how well three evaluation methods detect deficiencies in medical students' professionalism in two settings of an internal medicine clerkship. *Academic Medicine*. 2000;**75**:167-173.
[17] Hafferty FW, Levinson D. Moving Beyond Nostalgia and Motives: Towards a Complexity Science View of Medical Professionalism *Perspectives in Biology and Medicine*. 2008;**51**:599-615.
[18] Ginsburg S, Regehr G, Hatala R, et al. Context, Conflict, and Resolution: A New Conceptual Framework for Evaluating Professionalism. *Academic Medicine*. 2000;**75**:6-11.
[19] Ginsburg S, Regehr G, Lingard L. Basing evaluation of professionalism on observable behaviours: a cautionary tale. . *Academic Medicine*. 2004;**79**:S1-S4.
[20] Ginsburg S, Regehr G, Mylopoulos M. From behaviours to attributions: Further concerns regarding the evaluation of professionalism. *Medical Education*. 2009;**43**:414-425.
[21] Hilton SR, Slotnick HB. Proto-professionalism: how professionalisation occurs across the continuum of medical education.[see comment]. *Medical Education*. 2005;**39**:58-65.
[22] Quaintance JL, Arnold L, Thompson GS. Development of an Instrument fo Measure the Climate of Professionalism in a Clinical Teaching Environment. *Academic Medicine*. 2008;**83**:S5-S8.
[23] Haidet P, Kelly PA, Chou C. Characterizing the patient-centeredness of hidden curricula in medical schools: development and validation of a new measure. *Academic Medicine*. 2005;**80**:44-50.
[24] Rees CE, Knight LV. Viewpoint: The Trouble with Assessing Students' Professionalism: Theoretical Insights from Sociocognitive Psychology. *Academic Medicine*. 2007;**82**:46-50.
[25] Rees CE, Knight LV. Banning, detection, attribution and reaction: the role of assessors in constructing students' unprofessional behaviours. *Medical Education*. 2008;**42**:125-127.
[26] Stern D. *Mesuring medical professionalism*. New York Oxford University Press; 2006.
[27] Cruess R, Cruess S. Expectations and obligations: Professionalism and medicine's social contract with society *Perspect Biol Med*. 2008;**51**:579-598.

[28] Campbell EG, Regan S, Gruen RL, et al. Professionalism in Medicine: Results of a National Survey of Physicians. *Annals of Internal Medicine*. 2007;**147**:795-802.
[29] Walsh C, Abelson HT. Medical Professionalism: Crossing a Generational Divide *Perspectives in Biology and Medicine*. 2008;**51**:554-564.
[30] Holtman MC. A theoretical sketch of medical professionalism as a normative complex *Advances in Health Sciences Education*. 2008;**13**:233-245.
[31] Tsugawa Y, Tokuda Y, Ohbu S, et al. Professionalism Mini-Evaluation Exercise for medical residents in Japan: a pilot study. *Medical Education*. 2009;**43** 968-978.
[32] Ming-Jung Ho, Kun-Hsing Yu, Hirsh D, Tien-Shang Huang, Yang. P-C. Does One Size Fit All? - Building a Framework for Medical Professionalism. *Academic Medicine*. 2011 **In Press**.
[33] Hafferty FW, Castellani B. A sociological framing of medicine's modern-day professionalism movement *Medical Education*. 2009;**43**:826-828.
[34] Martimianakis MA, Maniate JM, Hodges BD. Sociological interpretations of professionalism. *Medical Education*. 2009;**43**:829-837.
[35] Kuper A, Reeves S, Levinson W. Qualitative research: An introduction to reading and appraising qualitative research. *BMJ*. 2008;**337:a288**:404-407.