

Professionalism Measures and Constructs

Administration and Scoring Guide

Project Team:

GEORGE NOWACEK, Ph.D., at the time of this study was the Director, Office of Medical Education at Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA

HERBERT M. SWICK, MD, is research professor, College of Health Professions and Biomedical Sciences, The University of Montana, Missoula, Montana; clinical associate professor, Department of Medicine, University of Washington School of Medicine, Seattle, Washington.

AMY V. BLUE, Ph.D., is the Assistant Provost for Education and Professor of Family Medicine at the Medical University of South Carolina (MUSC), Charleston, South Carolina, USA

SONIA CRANDALL, Ph.D., is a Professor of Family and Community Medicine at Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA.

RICHARD LUECHT, Ph.D., is a Professor of Educational Research Methodology at the University of North Carolina at Greensboro, Greensboro, North Carolina, USA.

SHEILA CHAUVIN, Ph.D., is the Director of the Academy for the Advancement of Educational Scholarship, and Professor in the Department of Medicine and Professor in the School of Public Health at Louisiana State University Health Sciences Center, New Orleans, Louisiana, USA.

J. CHARLES ELDRIDGE, Ph.D., is a Professor in the Department of Physiology and Pharmacology at Wake Forest University School of Medicine, Winston Salem, North Carolina, USA.

Acknowledgement

This study was supported through funding from the Stemmler Medical Education Research Fund (#IG33-0304).

TABLE OF CONTENTS

| | |
|---|-----------|
| * THE CONCEPTUAL BACKGROUND | 3 |
| * LIST OF RESOURCES | 4 |
| KNOWLEDGE-1: MEDICAL VIGNETTES | 4 |
| KNOWLEDGE-2: MULTIPLE CHOICE QUESTIONS | 4 |
| ATTITUDE-1: LIKERT-FORMAT STATEMENTS | 4 |
| ATTITUDE-2: SEMANTIC DIFFERENTIAL FORMAT..... | 4 |
| PEER ASSESSMENT-NOMINATIONS: | 4 |
| PEER ASSESSMENT-RATINGS OF ATTRIBUTES OF PROFESSIONALISM: | 4 |
| PROFESSIONALISM BEHAVIORS-FACULTY RATINGS IN SMALL GROUP SETTING: | 4 |
| PROFESSIONALISM BEHAVIORS-CLINICAL FACULTY RATINGS IN CLERKSHIP:..... | 4 |
| PROFESSIONALISM BEHAVIORS-COMMUNITY PHYSICIAN PRECEPTORS: | 4 |
| PROFESSIONALISM BEHAVIORS-COMMUNITY NURSING AND OFFICE STAFF: | 4 |
| BEHAVIORAL-OSCE FORMAT RATING OF SELF-REFLECTION: | 5 |
| BEHAVIORAL-OSCE FORMAT ETHICS CASE: | 5 |
| BEHAVIORAL-OSCE FORMAT PATIENT INTERACTION SKILLS (HUMANISM): | 5 |
| * EXPLANATION OF WHEN, HOW, AND THE ORDER IN WHICH TO USE EACH RESOURCE..... | 6 |
| 1) KNOWLEDGE AND ATTITUDE ASSESSMENTS OF PROFESSIONALISM | 6 |
| <i>Knowledge & Attitudes: Medical Vignettes, Multiple Choice Questions, Likert-format statements & Semantic Differential.....</i> | <i>6</i> |
| 2) PRE-CLINICAL ASSESSMENTS OF PROFESSIONALISM | 6 |
| <i>Peer Assessment: Nominations and Ratings.</i> | <i>6</i> |
| <i>Professionalism Behaviors-Faculty ratings in small group setting (Years 1 & 2):.....</i> | <i>6</i> |
| <i>Professionalism Behaviors: Community physician preceptors and Community nursing and office staff:</i> | <i>6</i> |
| 3) CLINICAL ASSESSMENTS OF PROFESSIONALISM..... | 7 |
| <i>Professionalism Behaviors-Clinical faculty ratings in clerkships (Year 3):</i> | <i>7</i> |
| <i>Peer Assessment: Nominations and Ratings.</i> | <i>7</i> |
| 4) CLINICAL SKILLS ASSESSMENT OF PROFESSIONALISM | 7 |
| <i>Behavioral-OSCE format assessments: Self-reflection exercise, Ethics case and Patient Interaction skills:</i> | <i>7</i> |
| * SCORING AND COMPUTING ANALYSIS CONSTRUCTS..... | 8 |
| KNOWLEDGE | 9 |
| <i>Measures used.....</i> | <i>9</i> |
| <i>Scoring</i> | <i>9</i> |
| <i>Computing Knowledge Construct.....</i> | <i>10</i> |
| ATTITUDE..... | 11 |
| <i>Measures used.....</i> | <i>11</i> |
| <i>Scoring</i> | <i>11</i> |
| <i>Computing Attitude Construct.....</i> | <i>12</i> |
| PRECLINICAL RATINGS..... | 13 |
| <i>Measures used.....</i> | <i>13</i> |
| <i>Scoring</i> | <i>13</i> |
| <i>Computing Preclinical Construct.....</i> | <i>14</i> |
| CLINICAL RATINGS | 15 |
| <i>Measures used.....</i> | <i>15</i> |
| <i>Scoring and Computing Clinical Construct</i> | <i>15</i> |
| CLINICAL SKILLS EXAM PERFORMANCE | 16 |
| <i>Measures used.....</i> | <i>16</i> |
| <i>Scoring</i> | <i>16</i> |
| <i>Computing Clinical Skills Construct</i> | <i>17</i> |
| * PRACTICAL IMPLEMENTATION ADVICE (MATERIALS NEEDED, LENGTH OF SESSION, FACULTY/FACILITATOR NEEDS, PREPARATION NEEDS, ETC.)..... | 18 |
| * HOW HAS IT BEEN SUCCESSFULLY DEPLOYED (INCLUDING COMMON PITFALLS, TIPS FOR SUCCESS, ETC.) | 18 |
| * WHAT ARE THE LIMITATIONS OF THE RESOURCE AND WHAT ARE YOUR IDEAS FOR IMPROVING/EXPANDING IT..... | 18 |

Professionalism Measures and Constructs

Administration and Scoring Guide

This manual has been developed to assist medical educators select and use one or more measures of professionalism for medical students. It reflects both theoretical and practical experience with administering the instruments to over 700 medical students in 8 classes at three medical schools over five years. While the longitudinal study administered all measures to one class over four years, it is recommended that selected measures be used to meet the expected outcomes of professionalism development at your medical school.

*** THE CONCEPTUAL BACKGROUND**

The 13 professionalism assessment measures were developed and evaluated during a 5-year project to develop a model of professionalism assessment of medical students. All measures were structured to measure five of the professionalism attributes identified by Swick (Swick H.M. (2000) Toward a normative definition of medical professionalism. *Academic Medicine*, 75(6), pp. 612-6.) and analysis of each measure maintained the separate attributes. The final path analysis model included all 13 measures in the educational timeline of preclinical, clinical and/or clinical skills constructs. The knowledge and attitude measures were administered during orientation week of medical school.

The path analysis model fit statistics confirmed that the obtained data fit the final model, but because the predictive power from one element of the model to the next was very low, it was concluded that there was not sufficient support for recommending the full model at this time. The positive results do support the efficacy of using the professionalism measures developed in the study in a 360-degree assessment model of medical student professionalism.

While all 13 measures included assessment of each of Swick's five attributes of professionalism, the student and faculty responses did not differentially distinguish behaviors or qualities of separate attributes. Even though data were combined across the attributes, it was concluded that future measures still need to include behaviors or items specific to a range of professionalism attributes. The assessment of individual component attributes was not supported at the undergraduate medical education level. However, these results do not support using a global 'professionalism' scale because it would not provide the same information as obtained in this study. Any measure of professionalism assessment needs to include multiple items that represent the component attributes of professionalism.

*** LIST OF RESOURCES**

Knowledge-1: Medical Vignettes

This measure of knowledge includes 15 vignettes that represent different aspects of professionalism in medical practice and are based on actual experiences. Selected vignettes from the AMIM Project Professionalism were used and additional ones were written to address the five attributes of this project.

Knowledge-2: Multiple Choice Questions

A second measure of knowledge of the professionalism attributes written in a standard multiple choice format. The items were written and revised by the project staff.

Attitude-1: Likert-format statements

The first measure of attitudes toward the five attributes of professionalism was a Likert-format measure with at least two items for each attribute of professionalism.

Attitude-2: Semantic Differential format

The second measure of attitudes toward the five attributes of professionalism used a Semantic Differential format, with each attribute as the focus concept. To insure that the entering students understood the meaning of the attribute, an abbreviated definition of the attribute was presented just above the statement of the attribute.

Peer Assessment-Nominations:

The literature on assessment of professionalism lists peer assessment as one of the most important approaches that should be used. Peer assessment is not a single methodology; it consists of at least two types of assessment. The nomination of peers who represent outstanding examples of the characteristic is easy, non-threatening, and reliable.

Peer Assessment-Ratings of attributes of professionalism:

The literature on assessment of professionalism lists peer assessment as one of the most important approaches that should be used. Peer assessment is not a single methodology; it consists of at least two types of assessment. The peer assessment included an on-line system in which students in their respective PBL groups were asked to rate each other on the individual attributes on a scale of very professional to very unprofessional

Professionalism Behaviors-Faculty ratings in small group setting:

A measure of student behaviors of professionalism used the specific behavioral scales developed for NBME funded project at the University of Miami. These behaviors were considered appropriate for the PBL small group setting. Faculty tutors were asked to rate each student on 11 specific behaviors and on the six attributes of professionalism on a scale of very professional to very unprofessional.

Professionalism Behaviors-Clinical faculty ratings in clerkship:

The same measure of student behaviors of professionalism developed for the PBL small group setting was used in the required Internal Medicine clerkship. This version of the measure was appended to the students' clinical evaluation by faculty attending and residents using an on-line computer system.

Professionalism Behaviors-Community physician preceptors:

The same measure of student behaviors of professionalism developed for the PBL small group setting was used in the community practice experience (CPE) setting. A paper form was used for the community faculty preceptors.

Professionalism Behaviors-Community nursing and office staff:

The same measure of student behaviors of professionalism developed for the PBL small group setting was used in the community practice experience (CPE) setting. A variation of the physician form was developed for the office nursing and administrative staff. It was suggested that student behavior in front of the physician might be different from that exhibited with nursing and office staff.

Behavioral-OSCE format rating of self-reflection:

The Clinical Practice Exam (CPX) is a multi-station exam in OSCE format with 12 to 16 cases. Each case includes 15 minutes with a standardized patient and 10 minutes after the encounter to complete specific tasks. For several stations, a self-reflection task was added to the post-encounter task.

Behavioral-OSCE format ethics case:

The Clinical Practice Exam (CPX) is a multi-station exam in OSCE format with 12 to 16 cases. Each case includes 15 minutes with a standardized patient and 10 minutes after the encounter to complete specific tasks. An ethics case used by Brown Medical school was inserted into the CPX beginning in 2002.

Behavioral-OSCE format patient interaction skills (humanism):

The Clinical Practice Exam (CPX) is a multi-station exam in OSCE format with 12 to 16 cases. Each case includes 15 minutes with a standardized patient and 10 minutes after the encounter to complete specific tasks. For every case, a common set of nine communications and relationship behaviors are rated by the standardized patient after the student encounter.

*** EXPLANATION OF WHEN, HOW, AND THE ORDER IN WHICH TO USE EACH RESOURCE**

The path analysis of the longitudinal data was optimized with the data clustered into four sets corresponding to phases of the medical curriculum: It is recommended that the measures be administered in sets to maximize the professionalism assessment information. It is not recommended to administer the measures individually.

The master copies of the measures are available in four separate files. The descriptions of when and how to administer the measures listed below are clustered according to the for file sets.

1) Knowledge and Attitude Assessments of Professionalism

Knowledge & Attitudes: Medical Vignettes, Multiple Choice Questions, Likert-format statements & Semantic Differential.

All four knowledge and attitude measures are included in single administration packet. Schedule about 40 minutes for administration time, but most students will finish in about 20 minutes.

Administered early in the first year, as early as orientation. Can also be administered during year 3 to measure changes, both increases and decreases, in professionalism knowledge and attitudes.

2) Pre-Clinical Assessments of Professionalism

Peer Assessment: Nominations and Ratings.

The peer ratings and nominations should be completed together. The assessment was computer administered in the study school because technology was used for all tests and assessments. A paper method was used in collaborating schools with similar results.

Peer assessment can and should be started in the first year of the curriculum. Peer ratings need to be based on scheduled interaction of students in small group or team settings. Any small group or team instruction can be used that lasts 1 hour or more over 6 or more sessions. (The length and duration of group meetings is based on antidotal impressions from the 5-year study but not empirically studied.)

Peer assessment can and should be used every time students participate in any group instruction. The peer assessment data is combined over multiple administrations to improve reliability. Peer assessment in clinical clerkships is possible and recommended when students rotate as a group for many or some of the services.

The peer assessment results were not feedback to students in the 5-year study. The decision to not provide feedback results was based, first, on student comments in focused group sessions. The other reason was that the 5-year study was designed as a purely assessment project. Feedback of results to students would have required structured sessions in the curriculum. These sessions would have required curriculum time and satisfy specific curriculum objectives.

Professionalism Behaviors-Faculty ratings in small group setting (Years 1 & 2):

Faculty were asked to complete the form for each student as an addition to the evaluation at the end of the term. Small group assignments were changed at each term; each student received four ratings in years 1 and 2. The evaluation forms were distributed and completed in paper form, one for each student.

Professionalism Behaviors: Community physician preceptors and Community nursing and office staff:

The community practice experience is completed during the summer between Years 1 & 2, and at the end of Year 2. The community practice physicians are mailed the professionalism behaviors form with the standard course evaluation forms for each student. For each student in the community practice office, one physician form and two nursing/staff forms were included. There were minor but important differences between the two forms. A separate self-addressed, stamped envelope was included with the professionalism forms to mail completed forms directly to the study staff.

3) Clinical Assessments of Professionalism

Professionalism Behaviors-Clinical faculty ratings in clerkships (Year 3):

Clinical faculty were asked to complete the form as an addition to the standard clinical evaluation form for each student. Each student received a variable number ratings depending on the number of clinical clerkships completed.

Peer Assessment: Nominations and Ratings.

The peer ratings and nominations should be completed together. The assessment was computer administered in the study school because technology was used for all tests and assessments. A paper method was used in collaborating schools with similar results.

Peer assessment can and should be started in the first year of the curriculum. Peer ratings need to be based on scheduled interaction of students in small group or team settings. Any small group or team instruction can be used that lasts 1 hour or more over 6 or more sessions. (The length and duration of group meetings is based on antidotal impressions from the 5-year study but not empirically studied.)

Peer assessment can and should be used every time students participate in any group instruction. The peer assessment data is combined over multiple administrations to improve reliability. Peer assessment in clinical clerkships is possible and recommended when students rotate as a group for many or some of the services.

The peer assessment results were not feedback to students in the 5-year study. The decision to not provide feedback results was based, first, on student comments in focused group sessions. The other reason was that the 5-year study was designed as a purely assessment project. Feedback of results to students would have required structured sessions in the curriculum. These sessions would have required curriculum time and satisfy specific curriculum objectives.

4) Clinical Skills Assessment of Professionalism

Behavioral-OSCE format assessments: Self-reflection exercise, Ethics case and Patient Interaction skills:

These assessments were included in, or added to, a clinical skills assessment (OSCE) completed by all students at the study school. These measures are not designed to be administered outside a comprehensive clinical skills assessment activity.

* SCORING AND COMPUTING ANALYSIS CONSTRUCTS

The analysis of the multivariate longitudinal data identified five constructs to represent the professionalism development of students in the 4-year undergraduate curriculum:

- Knowledge
- Attitude
- Preclinical Ratings
- Clinical Ratings
- Clinical Skills Examination Performance.

Table 1. Item Types and Scale Development Summary

| <i>Higher-Order Construct</i> | <i>Item Types</i> | <i>Number of Items/Tasks Developed</i> | <i>Final Items or Variables Used^a</i> | <i>Number of Score Scales</i> |
|-------------------------------|--|--|--|-------------------------------|
| Knowledge of Professionalism | MC & Vignette Items | 32 | 29 | 5 |
| Professionalism Attitudes | Likert Ratings & Semantic Differential | 26 | 26 | 5 |
| Preclinical Observations | Likert Ratings & Nominations | 321 | 73 | 4 |
| Clinical Observations | Likert Ratings & Nominations | 231 | 22 | 3 |
| Clinical Skills Exam | Protocol & Holistic Ratings | 72 | 72 | 3 |

^a Preclinical and clinical observational ratings on common items were aggregated over same-type sources (e.g. over all peers)

It is recommended to use the constructs for assessment using the multiple measures and scoring algorithms for the construct. It is not recommended to use a single measure since the reliability and validity data are based on the multiple measures, not on single measures.

KNOWLEDGE

Measures used

Knowledge-1: Medical Vignettes

Knowledge-2: Multiple Choice Questions

Scoring

Knowledge-1: Medical Vignettes

For items 1-6, the correct answer listed below is scored 1, incorrect 0

Vig1=A

Vig2=D

Vig3=D

Vig4=C

Vig5=A

Vig6=C

For items 7-15, score the item correct (1) if correct choice listed below is ranked 1 or 2

Vig7=B

Vig8=B

Vig9=A

Vig10=A

Vig11=D

Vig12=C

Vig13=A

Vig14=D

Vig15=D

Knowledge-2: Multiple Choice Questions

For each item, the correct answer listed below is scored 1, incorrect 0

MCQ1=A

MCQ2=A

MCQ3=A

MCQ4=A

MCQ5=C

MCQ6=C

MCQ7=D

MCQ8=C

MCQ9=A

MCQ10=A

MCQ11=A

MCQ12=A

MCQ13=A

MCQ14=B

MCQ15=B

MCQ16=A

MCQ17=C

Computing Knowledge Construct

For each factor, compute the mean of the variables listed (minimum valid responses)

Subordinating Self-interest

MCQ2, MCQ5, MCQ8, MCQ14, MCQ17, VIG14 (4)

Professional Responsibility

MCQ1, MCQ12, MCQ13, VIG6, VIG8, VIG10 (4)

Complexity & Uncertainty

MCQ4, MCQ10, MCQ15, VIG1, VIG5, VIG7, VIG13 (5)

Professional Commitment

MCQ3, MCQ7, MCQ9, VIG2, VIG3 (3)

Humanism

MCQ6, MCQ11, VIG9, VIG11, VIG15 (3)

Table 2. Knowledge Scales: Descriptive Statistics and Inter-Scale Correlations¹

| <i>Composite Measures</i> | <i>N</i> | <i>Mean</i> | <i>Std. Dev.</i> | <i>Correlations</i> | | | | |
|----------------------------------|----------|-------------|------------------|---------------------|----------|----------|----------|----------|
| | | | | <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> | <i>5</i> |
| <i>1:Knowledge-Service</i> | 212 | 3.939 | 1.220 | .456 | -.236 | -.168 | .481 | .375 |
| <i>2:Knowledge-Duty</i> | 212 | 4.311 | 1.079 | -.090 | .320 | -.217 | .073 | -.147 |
| <i>3:Knowledge-Uncertainty</i> | 212 | 3.458 | 1.099 | -.064 | -.069 | .317 | -.354 | .081 |
| <i>4:Knowledge-Personal Cost</i> | 212 | 2.387 | .989 | .189 | .024 | -.116 | .338 | -.349 |
| <i>5:Knowledge-Humanism</i> | 212 | 3.604 | .899 | .116 | -.038 | .021 | -.093 | .210 |

¹ The Cronbach "alpha" reliability coefficients are shown on the diagonal of the correlation matrix. The observed product-moment correlations are shown below the diagonal (lower triangle) and the disattenuated correlations are shown in the upper right triangle of the correlation matrix.

ATTITUDE

Measures used

Attitude-1: Likert-format statements

Attitude-2: Semantic Differential format

Scoring

Attitude-1: Likert-format statements

All items are coded SA=5 to SD=1, EXCEPT item #2 which is reversed coded SA=1 to SD=5

| | |
|-------------|-------------|
| ATT1 5 - 1 | ATT11 5 - 1 |
| ATT2 1 - 5 | ATT12 5 - 1 |
| ATT3 5 - 1 | ATT13 5 - 1 |
| ATT4 5 - 1 | ATT14 5 - 1 |
| ATT5 5 - 1 | ATT15 5 - 1 |
| ATT6 5 - 1 | ATT16 5 - 1 |
| ATT7 5 - 1 | ATT17 5 - 1 |
| ATT8 5 - 1 | ATT18 5 - 1 |
| ATT9 5 - 1 | ATT19 5 - 1 |
| ATT10 5 - 1 | ATT20 5 - 1 |
| | ATT21 5 - 1 |

Attitude-2: Semantic Differential format

Three bi-polar scales for each of 5 stimulus concepts are used: Good/Bad(G), Pleasant/Unpleasant(P), Awful/Nice(A)

The check marks on each scale are coded 7 to 1 when the check mark is in an interval delimited by colon (:). A check mark directly on a colon delimiter is coded to the higher value.

The scale variables by stimulus concept with direction of scale are:

| | |
|--|--|
| Ethical and Moral Standards EMG 7 - 1 EMP 7 - 1 EMA 1 - 7 | Humanistic Values HG 7 - 1 HP 7 - 1 HA 1 - 7 |
| Subordinate Self Interests SSG 7 - 1 SSP 7 - 1 SSA 1 - 7 | Self-reflection SRG 7 - 1 SRP 7 - 1 SRA 1 - 7 |
| Accountability AG 7 - 1 AP 7 - 1 AA 1 - 7 | |

Computing Attitude Construct

For each factor, compute the mean of variables listed (minimum valid responses)

Accountability

ATT10, ATT15, AG, AP, AA (3)

Humanistic Values

ATT3, ATT19, HG, HP, HA (3)

Self-Reflection

ATT2, ATT9, SRG, SRP, SRA (3)

Subordinate Self-interests

ATT4, ATT13, SSG, SSP SSA (3)

Ethics & Morals

ATT5, ATT14, ATT17, EMG, EMP, EMA (4)

Table 3. Attitude Scales: Descriptive Statistics and Inter-Scale Correlations¹

| <i>Composite Measures</i> | <i>N</i> | <i>Mean</i> | <i>Std. Dev.</i> | <i>Correlations</i> | | | | |
|---------------------------|----------|-------------|------------------|---------------------|----------|----------|----------|----------|
| | | | | <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> | <i>5</i> |
| <i>1:Attitudes-SSI</i> | 212 | 21.486 | 3.719 | .630 | .314 | .369 | .501 | .652 |
| <i>2:Attitudes-ISR</i> | 212 | 24.453 | 3.352 | .203 | .663 | .446 | .472 | .438 |
| <i>3:Attitudes-ECHV</i> | 212 | 28.080 | 2.508 | .230 | .285 | .616 | .489 | .608 |
| <i>4:Attitudes-ACCT</i> | 212 | 23.684 | 3.260 | .305 | .295 | .294 | .588 | .526 |
| <i>5:Attitudes-MEMS</i> | 212 | 31.198 | 3.021 | .427 | .294 | .394 | .333 | .681 |

¹ The Cronbach "alpha" reliability coefficients are shown on the diagonal of the correlation matrix. The observed product-moment correlations are shown below the diagonal (lower triangle) and the disattenuated correlations are shown in the upper right triangle of the correlation matrix.

PRECLINICAL RATINGS

Measures used

Peer Assessment-Nominations

Peer Assessment-Ratings of attributes of professionalism

Professionalism Behaviors-Faculty ratings in small group setting

Professionalism Behaviors-Community physician preceptors

Professionalism Behaviors-Community nursing and office staff

Scoring

Peer Assessment-Nominations

Student nominations were counted as the number of times the student was identified for one or more attributes.

The 'total nominations percent' (Totnomp) was calculated as sum of nominations by attribute over all phases, and then summed across attributes for a total number of nominations in all attributes in all phases.

The total number of nominations were tallied for each class and used as denominator for students in that class, resulting in a percent of nominations for a student in their class.

Peer Assessment-Ratings of attributes of professionalism

The ratings are the same for all attributes on a scale of 1=Extremely Unprofessional... To 7=Extremely Professional...

The peer assessment ratings were completed multiple times during years 1 and 2. Ratings were aggregated over all administrations within each attribute, and then computing the mean over all 6 professionalism attributes.

Professionalism Behaviors-Faculty ratings in small group setting

The faculty ratings of Behaviors of Professionalism consisted of two components: 9 behavior scales and 6 professionalism attribute scales.

Except for question #2, the 9 behavior scales are scored on a scale of 1=low and 7=high with descriptors anchoring the scale to be consistent with the content of the scale. For question #2, the response scale is recoded as follows: 4=7, 5,3=5, 6,2=3, 7,1=1

In the study, questions #1, #4, #5, #6, #7, #9 were averaged to a single index, bcp.qst

The 6 professionalism attributes scale responses were averaged to a single index bcp.rat.

Professionalism Behaviors-Community physician preceptors

The community physician ratings of Behaviors of Professionalism consisted of two components: 11 behavior scales and 6 professionalism attribute scales.

Except for question #3, the 11 behavior scales are scored on a scale of 1=low and 7=high with descriptors anchoring the scale to be consistent with the content of the scale. For question #3, the response scale is recoded as follows: 4=7, 5,3=5, 6,2=3, 7,1=1

In the study, questions #1, #2, #5, #6, #7, #8, #9 and all 6 attribute ratings were averaged to a single index, cp.md

Professionalism Behaviors-Community nursing and office staff

The community nursing and office staff ratings of Behaviors of Professionalism consisted of two components: 12 behavior scales and 6 professionalism attribute scales.

Except for question #3, the 11 behavior scales are scored on a scale of 1=low and 7=high with descriptors anchoring the scale to be consistent with the content of the scale. For question #3, the response scale is recoded as follows: 4=7, 5,3=5, 6,2=3, 7,1=1

In the study, questions #7, #8, #10, #11, #12,were averaged to a single index, cpstaff

Computing Preclinical Construct

Preclinical - Peer Ratings

Peer ratings

Preclinical - Peer Nominations

Peer nominations

Preclinical - CPE (Community Office Behavioral)

Mean of cp.md and cpstaff

Preclinical - BCSP (Faculty Small Group Behavioral)

Mean of bosp.qst and bosp.rat

Table 4. Preclinical Observations: Descriptive Statistics and Inter-Scale Correlations¹

| <i>Composite Measures</i> | <i>N</i> | <i>Mean</i> | <i>Std. Dev.</i> | <i>Correlations</i> | | | |
|-----------------------------------|----------|-------------|------------------|---------------------|----------|----------|----------|
| | | | | <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> |
| <i>1:Preclinical-Peer Ratings</i> | 212 | 4.052 | .279 | .895 | .246 | .232 | .258 |
| <i>2:Preclinical-Nominations</i> | 211 | .009 | .008 | .198 | .721 | .204 | .100 |
| <i>3:Preclinical-CPE</i> | 154 | 6.326 | .709 | .205 | .162 | .872 | .120 |
| <i>4:Preclinical-BCSP</i> | 217 | 5.817 | .611 | .222 | .077 | .102 | .827 |

¹ The Cronbach "alpha" reliability coefficients are shown on the diagonal of the correlation matrix. The observed product-moment correlations are shown below the diagonal (lower triangle) and the disattenuated correlations are shown in the upper right triangle of the correlation matrix.

CLINICAL RATINGS

Measures used

Peer Assessment-Nominations

Peer Assessment-Ratings of attributes of professionalism

Professionalism Behaviors-Clinical faculty ratings in clerkship

Scoring and Computing Clinical Construct

Peer Assessment-Nominations

The 'clinical nominations percent' (nominations) was calculated as sum of nominations across attributes for a total number of nominations in all attributes for phase 3 for each student.

The total number of nominations were tallied for each class and used as denominator for students in that class, resulting in a percent of nominations for a student in their class.

Peer Assessment-Ratings of attributes of professionalism

The ratings are the same for all attributes on a scale of 1=Extremely Unprofessional... To 7=Extremely Professional... The peer assessment ratings were completed during the medicine clerkship in year 3. Ratings were averaged over all 6 professionalism attributes for the index Peer Ratings

Professionalism Behaviors-Clinical faculty ratings in clerkship

The clinical faculty ratings of Behaviors of Professionalism consisted of two components: 11 behavior scales and 6 professionalism attribute scales.

Except for question #4, the 11 behavior scales are scored on a scale of 1=low and 7=high with descriptors anchoring the scale to be consistent with the content of the scale. For question #4, the response scale is recoded as follows: 4=7, 5,3=5, 6,2=3, 7,1=1. In the study, questions #1, #2, #4, #5, #6, #7, #8, #9 #10 #11 and all 6 attribute ratings were averaged to a single index, Faculty Ratings.

Table 5. Clinical Observations: Descriptive Statistics and Inter-Scale Correlations¹

| <i>Composite Measures</i> | <i>N</i> | <i>Mean</i> | <i>Std. Dev.</i> | <i>Correlations</i> | | |
|------------------------------------|----------|-------------|------------------|---------------------|----------|----------|
| | | | | <i>1</i> | <i>2</i> | <i>3</i> |
| <i>1: Clinical-Faculty Ratings</i> | 208 | 6.497 | .391 | .958 | .195 | .238 |
| <i>2: Clinical-Peer Ratings</i> | 208 | 4.174 | .444 | .186 | .948 | .385 |
| <i>3: Clinical-Nominations</i> | 202 | .986 | .939 | .202 | .325 | .753 |

¹ The Cronbach "alpha" reliability coefficients are shown on the diagonal of the correlation matrix. The observed product-moment correlations are shown below the diagonal (lower triangle) and the disattenuated correlations are shown in the upper right triangle of the correlation matrix.

CLINICAL SKILLS EXAM PERFORMANCE

Measures used

OSCE format rating of self-reflection

OSCE format ethics case

OSCE format patient interaction skills (humanism)

Scoring

The scoring rubrics for the self-reflection questions presented below were developed from a content analysis of responses from about 300 administrations of the measure.

OSCE format rating of self-reflection

| | | |
|---|---|--|
| <p>Question #1 <i>[How do you think you did with this case?]</i> 0=No judgement 1=Not well, struggled 2=Not sure 3=Fine, Fair 4=Ok, good, well 5=Great 8=No Answer</p> | <p>Question #2 <i>[If you could repeat this case, would you do anything differently?]</i> 1=Nothing different 2=Interview process, structure 3=Engage more with patient 4=Change self, behavior 5=More, different Hx, Pe 6=Uncertain 8=No Answer</p> | <p>Question #3 <i>[Was there anything you learned about yourself from this case?]</i> 1=Nothing 2=Medical knowledge 3=Medical decision making 4=Career interests 5=My interview style 6=My emotions, affect 7=Incidental, peripheral 8=No Answer</p> |
| | | <p>Q#3 Response Content 1=No reflection 2=Process 3=Knowledge, skill, DD 4= Self 5=Hx, PE</p> |

Judgment of Self Reflection?

{ Subjective judgment from review of each case responses }

0=No, restate events

1=Yes, self reflection response

OSCE format ethics case

| | | |
|---|---|--|
| Number of issues # of valid ethics issues 0 to 3 | Ethics Discussion 0=Insufficient, none 1=Minimal discussion 2=OK discussion 3=Excellent discussion | SP Questions {5 questions} 5 =Excellent job, everything done right 4 = Good Job, most things done 3 = Acceptable work, no major problems 2 = Marginal, missed some major issues 1 = Very problematic, missed major points |
|---|---|--|

OSCE format patient interaction skills (humanism)

6 Communications Questions, 3 Relationship Questions

- 1=Poor
- 2=Fair
- 3=Good
- 4=Very Good
- 5=Excellent

Computing Clinical Skills Construct

Standardized Patient Ratings (CPX - SP Performance)

Mean of 5 SP Questions with at least 2 valid ratings

Ethics Case (CPX – Ethical Problem)

Mean of Number of issues and Ethics Discussion,

Self-Reflection (CPX – Self Reflection)

Mean of judgment of self-reflection (0,1) from 3 cases with at least 2 valid responses.

Table 6. CPX Professionalism Indicators: Descriptive Statistics and Inter-Scale Correlations¹

| <i>Composite Measures</i> | <i>N</i> | <i>Mean</i> | <i>Std. Dev.</i> | <i>Correlations</i> | | |
|---------------------------|----------|-------------|------------------|---------------------|----------|----------|
| | | | | <i>1</i> | <i>2</i> | <i>3</i> |
| 1:CPX-SP Performance | 197 | 3.944 | .478 | .814 | .001 | .075 |
| 2:CPX-Ethical Problem ID | 197 | 1.751 | .819 | .001 | .684 | .076 |
| 3:CPX-Self Reflection | 196 | .373 | .346 | .046 | .043 | .463 |

¹ The Cronbach “alpha” reliability coefficients are shown on the diagonal of the correlation matrix. The observed product-moment correlations are shown below the diagonal (lower triangle) and the disattenuated correlations are shown in the upper right triangle of the correlation matrix.

*** PRACTICAL IMPLEMENTATION ADVICE (MATERIALS NEEDED, LENGTH OF SESSION, FACULTY/FACILITATOR NEEDS, PREPARATION NEEDS, ETC.)**

- Minimal faculty preparation. Instructions for administration are provided with the measures.
- Institutional, curriculum commitment to professionalism assessment necessary. Cannot be individual faculty or course effort.
- Administration of the measures does not require curriculum changes in instruction in professionalism.
- Administration of measures will provide detailed information about individual student professionalism development for curriculum evaluation or quantification of student status.

*** HOW HAS IT BEEN SUCCESSFULLY DEPLOYED (INCLUDING COMMON PITFALLS, TIPS FOR SUCCESS, ETC.)**

The measures have been used with five incoming classes in the study school and six classes in collaborating medical schools. There has not been any significant opposition to the administration of the measures nor the content of the measures in any of the eleven classes. The project staff worked with individual course directors to enlist cooperation for administration of the measures. For student participation, cooperative participation was obtained by beginning administration of the measures during orientation that included an overall description of the 4-year study. They were also told that this was a study of the assessment of professionalism and that there would not be any feedback from any of the measures.

*** WHAT ARE THE LIMITATIONS OF THE RESOURCE AND WHAT ARE YOUR IDEAS FOR IMPROVING/EXPANDING IT.**

With the exception of the clinical skills measures, the measures can be administered within the framework of the assessment activities of the traditional medical school curriculum. They can be appended to the ongoing assessments in small group activities, clinical evaluations and other assessments without overburdening the system. The clinical skills measures are designed to be included in an OSCE or other clinical skills assessment exercise.