Communication Assessment Using the Common Ground Instrument: Psychometric Properties

Forrest Lang, MD; Ronald McCord, MD; Leo Harvill, PhD; Delia S. Anderson, MS

Background and Objectives: *Recent guidelines from the Association of American Medical Colleges and from the Accreditation Council for Graduate Medical Education strongly suggest that communications teaching and assessment be part of medical education at all levels. This study's objective was to validate an instrument to assess communications skills. This instrument, Common Ground, is linked to the core, generic communication skills emphasized by the consensus statements of Toronto and Kalamazoo.* Methods: *A total of 100 medical students were recruited from two medical schools and tested with four-station, communications-focused objective structured clinical examinations. Using Common Ground, trained raters performed checklist and global rating assessments. Experts globally assessed 20 representative interviews.* Results: *Inter-rater reliability for Common Ground was 0.85 for the overall global ratings and 0.92 for the overall checklist assessment. Generalizability coefficient was 0.80 for 50 minutes of testing. The correlation between the ratings of trained raters and a panel of communication experts was 0.84.* Conclusions: *The Common Ground assessment instrument assesses core communication skills with sufficient reliability, validity, and generalizablity to make decisions on medical students' performance.*

(Fam Med 2004;36(3):189-98.)

Assessment of the communication skills of medical students and residents is now recognized as a necessary component of medical education. The Association of American Medical Colleges (AAMC) recently published recommendations that all medical schools should include communication instruction, practice, and assessment as a part of a clinical competencies curriculum.[1] The National Board of Medical Examiners (NBME) is developing an objective structured clinical examination (OSCE) that includes assessment of communication skills as part of its United States Medical Licensure Examination (USMLE), and the Accreditation Council for Graduate Medical Education (ACGME) directed all residencies to teach and assess communications as a core clinical skill.[2] Recent consensus statements by international experts identify a discrete set of core communication skills[3] with compelling evidence of their positive effect on medical communications.[4] If all medical schools and residencies teach and assess communication skills as recommended, how should this assessment be performed when it is linked with high-stakes outcomes, such as passing courses, completing graduation requirements, and securing medical licensure?

The current emphasis by experts in the field expands the communications domain thought necessary for effective communications. For example, both consensus statements include an emphasis on those patient-centered skills with which family medicine research has been so closely associated. New assessment instruments are needed to evaluate communication skills, especially patient-centered skills, such as establishing focus, active listening, and the use of communication to reach common ground.[5-7]

Attempts to assess communication skills have proven challenging on a number of fronts. Some of the problems include a high case-to-case variation in the skills to be assessed and the way in which learning demonstrates those skills.[8-10] Such variation led one author to state that "No universal set of communication skills exists that can be assessed."[9] Indeed, a number of studies using a variety of assessment techniques demonstrate inadequate inter-rater agreement and poor generalizability,[11,12] particularly when using faculty to assess learners' communication skills.

From the Department of Family Medicine (Drs Lang and McCord [deceased]) and the Section of Medical Education (Dr Harvill), East Tennessee State University; and the Department of Medicine, Tulane University (Ms Anderson).

Certain assessment methods that have demonstrated reliability and generalizability have built their assessment on a more limited domain of communication assessment,[8,13] some appearing to measure a single dimension of assessment, such as overall satisfaction. Although several of the efforts to assess a wider domain of communication skills in a high-stakes situation have demonstrated reasonable success, such successes are limited by their need to use expert faculty to perform the assessment[14] or by the unrealistic requirements for time of observation (2.5 hours or at times much longer).[15]

To address these concerns, we developed a new communications assessment instrument, Common Ground, built on the template of the Toronto and Kalamazoo Consensus Statements. The research reported here explores whether Common Ground can reliably, validly, and efficiently (acceptable generalizability) assess learners' communication skills during performance of a patient-centered interview.

Methods

Case Scenarios

We began by identifying the core communication skills specified by the Toronto and Kalamazoo Consensus Statements.[3,4] We created a set of standardized patient (SP) scenarios, each built to allow multiple assessments of these core communication skills: rapport building,[16] information management, agenda setting,[7] active listening,[6] addressing feelings,[17] and negotiations to reach common ground.[18]

To limit case-specific variation, the cases involved a variety of common office-based problems ranging from sore throat to abdominal pain to cervical arthritis. To take advantage of the merits of both checklists and global assessments,[19,20] we developed an assessment instrument called Common Ground that both checked the performance of skills in response to the built-in opportunities/clues and also developed a global, criteria-based rating guide for each skill and for the interview as a whole. For cases that did not involve an opportunity to develop a plan incorporating shared decisions between clinicians and patients, the Common Ground assessment section was left off the assessment instrument.

Subjects

Faculty and students at two medical schools participated in testing the communication assessment system. One of these schools provided students with extensive experience with SPs in clinical teaching and assessment but provided little (limited to 2 hours of lecture and demonstration during the first-year orientation week) communication instruction—the minimal interviewing curriculum. The other school uses SPs working with experienced faculty as part of eight 2-hour communication workshops during the second year. Skills were reinforced at the second school during the third year since students are observed and critiqued three times when interviewing SPs during their family medicine clerkship.

Two cohorts of 25 students were recruited from each school. One cohort of students was recruited from the entering classes of each school in August 1998. A second cohort was recruited from each school from students about to become juniors. This second pair of cohorts was used to assess communication skills at the interface between preclinical and clinical curricula and was tested an additional time at the end of the junior year. Students were compensated for their time.

During each assessment period, students conducted videotaped interviews with four SPs each for 10 minutes.[21] Two SPs cases were used in common in all three assessments.

SPs and Raters

SP training occurred prior to testing and included an assessment of the delivery and the timing of presentation of clues embedded in the cases. SPs were certified for use when they were able to deliver 85% of the clues accurately and on time.

Raters for the project were recruited and trained using a guidebook developed as part of this project. Four raters responded to a newspaper ad for individuals with at least 2 years of college education. Our *a priori* accuracy goal (agreement of ratings with the consensus rating of two of the authors) to begin official rating was set at 80%. One of the raters failed to reach this level of accuracy and did not continue. One trained rater's schedule changed, limiting her availability to perform ratings. Two of the raters, a college student (rater #1) and a retired teacher (rater #2), performed the majority of study ratings.

Rating Process

The case scenarios were randomized so that students from both schools performed a variety of case scenarios. Raters were blind to the school and year of the student in each interview. The data obtained from trained raters on each videotaped interview included percentage scores for each of the six core skills. The percentage scores were determined by dividing the number of points obtained for a skill by the maximum number of points possible. An overall examination percentage score was also calculated from the unweighted percentage scores of the six skills. In addition, raters recorded a global rating of each skill and of the overall interview based on the criteria-based global rating scale.

Psychometric Assessment

A variety of methods were used to assess the psychometric qualities of Common Ground. To assess internal consistency, each rater's percentage scores were compared with their global ratings for each skill and for the overall interview. To assess test-retest reliabil-

ity (intra-rater), rater #1 repeated the scoring of 12 randomly selected interviews while rater #2 repeated the scoring of 10 randomly selected interviews with approximately 15 weeks intervening. To assess inter-rater reliability, both raters scored the same set of 10 interviews; Pearson correlation coefficients were calculated between the ratings and between the percentage scores. Generalizability coefficients were used to determine if the assessments are consistent across a number of varied clinical cases. The generalizability coefficient (rho squared) represents the reliability of an assessment when one student is compared to the performance of others while the dependability coefficient (phi) is the reliability of an assessment when a student's performance is compared to an absolute standard rather than being compared to the performance of others.

Regarding construct validity, we compared assessments of the randomly selected first-year cohort to the assessments of the randomly selected third- or fourth-year students with the expectation that performance should improve with intensive instruction. The final validity study was the comparison of the ratings of the six skills and overall performance by five patient-centered communication experts with the same ratings and the percentage scores from the raters. The ratings and percentage scores from the raters were compared with the mean rating of the five experts. Agreement between the two groups represents a measure of concurrent validity and would indicate that trained raters can provide assessments similar to experts. The data from the two groups based on the assessment of 20 interviews was compared with Pearson correlation coefficients.

Five faculty experts from across the United States, each of whom teach in national faculty development communications courses, viewed 20 interviews selected by the authors. The interviews were selected by purposive sampling rather than randomly selected to assure that the interviews included the full range of performance from those receiving the higher, the lower, and the mid-range scores. The experts globally rated each core skill and the overall interview.

Results

Checklist Versus Global Rating

Intra-rater consistency between the case percentage scores and the global ratings showed excellent agreement for each rater. The Pearson correlation coefficients were 0.95 and 0.91, respectively.

Intra-rater Reliability

The agreement between the two assessments of the same interview for specific items was excellent for some of the skills but less so for others (Table 1). For rater #1, the overall case rating had an intra-rater correlation of 0.63, and the overall case percentage scores had a correlation of 0.69. For rater #2, the overall case rating had an intra-rater correlation of 0.87, and the overall case percentage scores had a correlation of 0.78. Rater #1 had higher individual skill correlations while rater #2 had higher overall case correlations.

Inter-rater Reliability

Correlations for the various global ratings were: rapport building (0.49), information management (0.63), agenda setting (0.79), active listening (0.86), addresses feelings (0.97), and overall case rating (0.85). The correlations for the various checklist percentage scores were: rapport building (0.60), information management (0.83), agenda setting (0.69), active listening (0.86), addresses feelings (0.95), and overall case (0.92). We did not compute correlations for the "reaching common ground" global rating or percentage score because half of the interviews assessed did not include this component.

Generalizability

Generalizability coefficients are presented for the rising juniors in Table 2. Similar analyses with similar results were completed on the groups of entering students and students at the end of the third year but are not reported here. When using all four cases in the analysis with the dependent variable of overall case percentage score, it was determined that five cases were needed to achieve a generalizability coefficient of 0.80 or

Table 1

Intra-rater Reliability—Pearson Correlation Coefficients Between Repeated Scoring of Randomly Selected Interview for Two Raters

*Score—Global Rating**	*(n=12) Rater #1*	*(n=10) Rater #2*
Rapport building	-0.12	0.62
Information management	0.76	0.53
Agenda setting	0.79	0.38
Active listening	0.90	0.75
Addresses feelings	0.67	0.58
Common ground	0.07	0.23
Overall case	0.63	0.87
*Score—Checklist %***		
Rapport building	0.09	0.53
Information management	0.50	0.46
Agenda setting	0.88	0.75
Active listening	0.80	0.95
Addresses feelings	0.83	0.33
Common ground	0.67	0.60
Overall case	0.69	0.78

* Rating refers to the global rating of that skill (see Appendix B for Global Rating Guide)

** % score refers to the performance on the behavioral checklist as a percent of the maximum possible score (See Appendix A for Common Ground Checklist)

Table 2

Generalizability Coefficients If Different Numbers of Stations/Cases Are Used—Rising Juniors

Rapport building %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 * 0.31		0.48	0.58	0.65	0.70	0.73	0.78	**0.82**
ϕ** 0.27		0.43	0.53	0.60	0.65	0.69	0.75	0.79
Information management %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.34		0.51	0.61	0.68	0.72	0.76	**0.81**	0.84
ϕ 0.33		0.49	0.59	0.66	0.71	0.74	**0.80**	0.83
Agenda setting %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.23		0.37	0.47	0.55	0.60	0.64	0.71	0.75
ϕ 0.22		0.36	0.46	0.53	0.59	0.63	0.70	0.74
Active listening %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.36		0.53	0.62	0.69	0.74	0.77	**0.82**	0.85
ϕ 0.26		0.41	0.51	0.58	0.63	0.67	0.73	0.78
Addresses feelings %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.27		0.42	0.52	0.59	0.64	0.68	0.74	0.78
ϕ 0.24		0.38	0.48	0.55	0.61	0.65	0.71	0.76
Common ground %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.13		0.24	0.32	0.38	0.44	0.48	0.55	0.61
ϕ 0.13		0.24	0.32	0.38	0.43	0.48	0.55	0.61
Overall case %								
G Study	If n of							
n of cases=1	cases =	2	3	4	5	6	8	10
ρ^2 0.45		0.62	0.71	0.77	**0.80**	0.83	0.87	0.89
ϕ 0.38		0.56	0.65	0.71	0.76	0.79	**0.83**	0.86

* ρ^2 represents the reliability of an assessment when one student is compared to the performance of others.

** ϕ represents the reliability of an assessment when a student's performance is compared to an absolute. Bolded figures indicate the smallest number of cases for which a coefficient of .80 is achieved.

greater (0.80) and that eight cases would provide a dependability coefficient of 0.80 or greater (0.83).

Validity

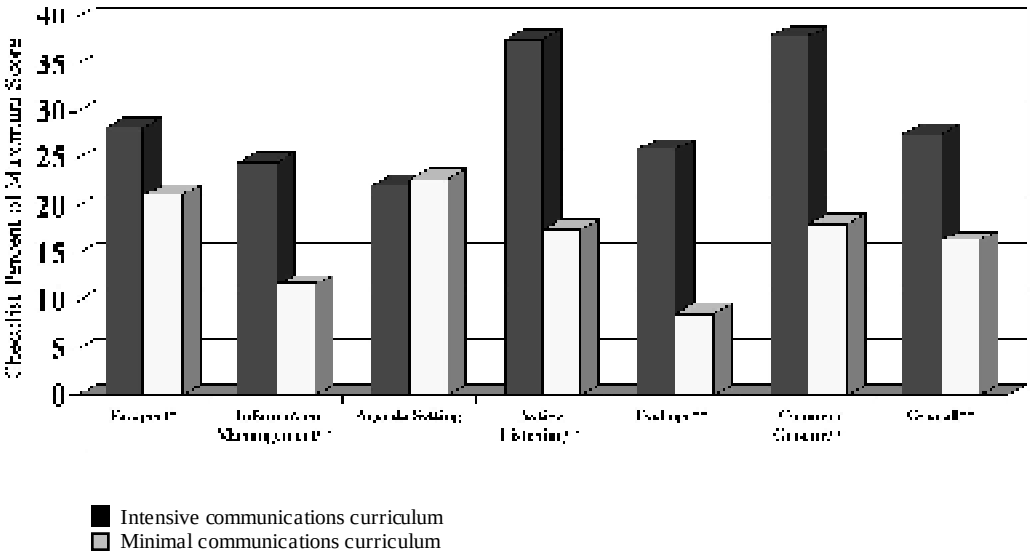
Regarding construct validity, during the first week of medical school, students from the minimal interviewing curriculum school scored slightly higher (but statistically significant) than students at the interview-intensive curriculum school scored. Later, in both assessments of clinical year students at the interview-intensive curriculum school, students at the interview-intensive school outperformed students at entry to either school on five of six of the core skills and on the overall interview percentage scores and on all ratings (Figure 1). At the minimal interviewing curriculum school, there were no statistically significant differences on any of the core skills or on the overall interview assessment between entry-level students and those tested at either time in their clinical years.

Regarding concurrent validity, the findings are presented in Table 3. There was good agreement between the expert mean rating and the rater ratings and per-

Figure 1

Third-year Students' Communication Performance

Intensive Communications Curriculum School Versus Minimal Communications School



* *P*<.025
** *P*<.001

centage scores for the skills of information management, active listening, and addresses feelings, as well as for the overall performance. The correlations for rapport building, agenda setting, and reaching common ground were not as good. The correlation for the overall performance between the expert mean rating and the rater rating was 0.84 and between the expert mean rating and the rater percentage score was 0.83.

Discussion

The rater-to-independent-expert validity check of the Common Ground instrument provides evidence that the Common Ground instrument checklist and global assessments capture impor-

Table 3

Concurrent Validity—Agreement Between Trained Raters and Experts

Pearson correlation coefficients (n=20)

	Trained rater RB rating	Trained rater RB %
Expert rapport building (RB) mean rating	0.55	0.62
	Trained rater IM rating	Trained rater IM %
Expert information management (IM) mean rating	0.81	0.90
	Trained rater AS rating	Trained rater AS %
Expert agenda setting (AS) mean rating	0.57	0.37
	Trained rater AL rating	Trained rater AL %
Expert active listening (AL) mean rating	0.84	0.85
	Trained rater AF rating	Trained rater AF %
Expert addresses feelings (AF) mean rating	0.82	0.81
	Trained rater CG rating	Trained rater CG %
Expert common ground (CG) mean rating	0.33	0.62
	Trained rater overall rating	Trained rater overall %
Expert overall mean rating	0.84	0.83

tant elements of a competent patient-centered interview. In Boon and Stewart's 1998 review of 44 assessment instruments, only the Maastricht History Taking and Advise Checklist had comparable convergent validity with global expert ratings.[22]

Even though Common Ground assesses a more robust domain of communication skills than the Educational Commission for Foreign Medical Graduates and NBME, the number of cases needed to assure a stable and representative assessment with Common Ground is less than with other instruments in the literature.[8,14,15] We attribute this efficiency in assessment to (1) the choice of routine office cases as opposed to selecting cases requiring case-specific communications skills and (2) scenario development that provides multiple opportunities to demonstrate each of the core skills within each case. The more opportunity that a person has to demonstrate skills, or the lack of them, the more stable the profile of performance becomes. We are confident that when an interviewer fails to respond to multiple stimuli in three to four interviews, skills in that particular area are truly deficient and require remediation. There is, however, a downside to incorporating multiple opportunities to demonstrate each of the skills in each case. The multiple clues may eventually prompt an interviewer's response (thus inflating the performance assessment), whereas a longer set of interviews with fewer clues might not gain the attention of the same interviewer.

The data in this study support the premise that there is a set of generic or core communication skills[3,4] that have applicability to a wide range of everyday encounters and that these skills can be reliably assessed. This study should encourage programs to use communication scenarios and assessment instruments that are tied to existing consensus statements and evidence-based communications research.

Limitations

This project is limited in several ways. The generalizability of assessment is limited to communication skills with patients who have everyday symptoms or problems. It is not possible to generalize to handling special, challenging situations like breaking bad news, counseling with patients with alcoholism, or interviewing a psychotic patient. Handling such situations is important and should be a part of every communications curriculum but is not part of the communications skills assessed by Common Ground.

Certain areas of skill assessment appear to need additional work. The rapport-building correlations between raters and experts were relatively weak. We are in the process of assessing the influence of verbal rapport-building statements ("positive speak" on the assessment instrument) and nonverbal skills (paralinguistics and proximics) on rater and expert assessment. In addition, assessing Common Ground skills needs further work. We and other researchers are in the process of clarifying these skills and their impact on successful interviewing.

While the positive improvement in skill performance after instruction is reassuring, it is not possible to say how much training is required to approach a maximum benefit. It is possible that identical results would have occurred with half the workshops/practice, but it is also possible that the behavioral impact would have doubled if twice the practice time had occurred. More study is necessary to establish the ideal amount and optimal timing of communication instruction.

Conclusions

The Common Ground Communications Assessment Instrument provides a reliable and valid assessment of patient-centered communications skills for everyday office visits, consistent with the expectations of the Toronto and Kalamazoo Consensus Statements. When paired with cases designed to provide multiple opportunities to demonstrate these skills, it becomes feasible to assess these complex skills in high-stakes examinations.

Acknowledgments: This research was supported by a grant from the Stemmler Medical Education Research Fund of the National Board of Medical Examiners.

Ron McCord provided energy and vision to this project until his untimely death in the summer of 2001. He and his contributions are missed.

Corresponding Author: Address correspondence to Dr Lang, East Tennessee State University, Department of Family Medicine, Box 70621, Johnson City, TN 37614. 423-439-5828. Fax: 423-439-5795. lang@mail.etsu.edu.

References

1. Association of American Medical Colleges. Contemporary issues in medicine: communication in medicine. Washington, DC: Association of American Medical Colleges, 1999.
2. Accreditation Council for Graduate Medical Education. Outcome Project. www.acgme.org/outcome.
3. Participants in the Bayer-Fetzer Conference on Physician-Patient Communication in Medical Education. Essential elements of communication in medical encounters: the Kalamazoo consensus statement. Acad Med 2001;76:390-3.
4. Simpson M, Buckman R, Stewart M, et al. Doctor-patient communication: the Toronto consensus statement. BMJ 1991;303:1385-7.
5. Stewart M, Brown JB, Weston WW, McWhinney IR, McWilliam CL, Freeman TR. Patient-centered medicine. London: Sage Publications, 1995.
6. Lang F, Floyd MR, Beine KL. Clues to patients' explanations and concerns about their illnesses. A call for active listening. Arch Fam Med 2000;9:222-7.
7. Mauksch L, Hillenburg L, Robins L. The establishing focus protocol: training for collaborative agenda setting and time management in the medical interview. Families, Systems, and Health 2001;19:147-57.
8. Boulet JR, Ben-David MF, Ziv A, et al. High-stakes examinations: what do we know about measurement? Using standardized patients to assess the interpersonal skills of physicians. Acad Med 1998;73(10 suppl):S94-S96.
9. Hodges B, Turnbull F, Cohen R, Bienenstock A, Norman G. Evaluating communication skills in the objective structured clinical examination format: reliability and generalizability. Med Educ 1996;30:38-43.
10. Kroboth FJ, Hanusa BH, Parker S, et al. The inter-rater reliability and internal consistency of a clinical evaluation exercise. J Gen Intern Med 1992;7:174-9.

11. Stillman P, Swanson D, Regan MB, et al. Assessment of clinical skills of residents utilizing standardized patients: a follow-up study and recommendations for application. Ann Intern Med 1991;114(5):393-401.
12. Kalet A, Earp JA, Kowlowitz V. How well do faculty evaluate the interviewing skills of medical students? J Gen Intern Med 1992;7:499-505.
13. Cohen DS, Colliver JA, Marcy MS, Fried ED, Swartz MH. Psychometric properties of a standardized patient checklist and rating scale form used to assess interpersonal and communication skills. Acad Med 1996; 71:S87-S89
14. Brown JB, Handfield R, Rainsberry P, Brailovsy C. Certification examination of the College of Family Physicians of Canada. Can Fam Physician 1996;42:1539-48.
15. Van Thiel J, Kraan F, Van der Vleuten CPM. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist. Med Educ 1991;25:224-9.
16. Thom DH, Campbell B. Patient-physician trust: an exploratory study. J Fam Pract 1997;44:169-76.
17. Platt FW, Platt CM. Empathy: a miracle or nothing at all? JCOM 1998; 5:30-3.
18. Taylor TR. Understanding the choices that patients make. J Am Board Fam Pract 2000;13:124-33.
19. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. Acad Med 1998;73(9):993-7.
20. Swartz M, Colliver J, Bardes C, Charon R, Fried E, Moroff S. Global ratings of videotaped performance versus global ratings of actions recorded on checklists: a criterion for performance assessment with standardized patients. Acad Med 1999;74(9):1028-32.
21. Van Mourik TGC, van der Vleuten CPM. Authentic assessment of interviewing and counseling skills: effect of testing time per station on generalizability and validity. Teach Learn Med 1995;7:155-62.
22. Boon H, Stewart M. Patient-physician communications assessment instruments: 1986 to 1996 in review. Patient Educ Couns 1998;35:161-76.

Appendix A

Common Ground Checklist/Global Assessment

Interviewer _____ Faculty _____ Date _____ S. P. (Generic)

1. Rapport

No	1	2	3	4	5	
O	O					Initial introduction to patient
O	O	O	O	O	O	Explicit "Positive Speak"
O	O	O	O	O	O	Explicit caring/commitment
O	O	O	O	O	O	Verbal interruption
O	O	O	O	O	O	Negative talk (implied or explicit)

Nonverbal Interest	-2 Strong Negative	-1 Negative	0 Neutral	+1 Positive	+2 Strong Positive
Body [position, etc]	O	O	O	O	O
Eye contact	O	O	O	O	O
Voice Qualities	O	O	O	O	O

1	2	3	4	5	
O	O	O	O	O	Overall Rapport

2. Information Management

0	1	2	3	4	
O	O	O	O	O	For the first interviewer's questions record the number of closed ended questions.

If 10 or 9 – record as 0 O _____
If 8 or 7 – record as 1 O
If 6 or 5 – record as 2
If 4 or 3 – record as 3
If 2 or 1 – record as 4

0	1	2	3	4	
O	O	O	O	O	Performs summary with 5 or more items.

1	2	3	4	5	
O	O	O	O	O	Overall Information Management

3. Eliciting all Agenda Items

No	1	2.	3	
O	O	O	O	Agenda elicited (Step 1) "What brings you in?" "How can I help?"
O	O			Early full exploration of patient's agenda (Step 2) "That's it?"
O	O		O	Checks for additional agenda later in the interview.

1	2	3	4	
O	O	O	O	Overall Agenda

4. Active Listening for Full Understanding of Ideas, Concerns, and Expectations

No	Yes	N/A	PT's clues or statements needing follow up.
O	O	O	Clue #1-
O	O	O	Clue #2-
O	O	O	Clue #3-
O	O	O	Clue #4-
O	O	O	Clue #5-
O	O	O	Clue #6-

0	1	2	3	4	
O	O	O	O	O	Asks for clarification about patients' ideas, concerns, expectations.

1	2	3	4	5	
O	O	O	O	O	Overall Active Listening

Observations and Comments

(continued on next page)

Appendix A

(continued)

Interviewer _____ Faculty _____ Date _____ S. P. (Generic)

5. Addressing Feelings with Patient

No	Yes	N/A	Pt. stated or implied feelings/concerning follow up
O	O	O	**Feeling #1-**
O	O	O	**Feeling #2-**
O	O	O	**Feeling #3-**
O	O	O	**Feeling #4-**

0	1	2	3	4	Evidence of ability to deal with feelings
O	O	O	O	O	

1	2	3	4	5	**Overall Addressing Feelings**
O	O	O	O	O	

6. Reaching Common Ground
Checks Feasibility and Understanding

No	1	2	3	
O	O	O	O	Checks for agreement/feasibility
O	O	O	O	Checks for understanding
O	O	O	O	Defines mutual responsibility

0	1	2	3	4	Less Effective Strategies
O	O	O	O	O	Routine prescription, personal preferences, methods/modalities or medical/disease-centered recommendations
O	O	O	O	O	**More Effective Strategies** Assessing readiness to change, exploring patient's ideas, knowledge, expectations, making patient-centered recommendations, reframing, brainstorming, performing decision analysis, giving options, and priorities

1	2	3	4	5	**Overall Reaching Common Ground**
O	O	O	O	O	

7. Global Interview Performance

1	2	3	4	5	**Overall Global Interview**
O	O	O	O	O	

Observations and Comments

Appendix B

Global Rating Guide of Core, Common Ground Interview Skills

Rapport Building—Global Criteria

5. Demonstrates rapport-building skills such that most patients would subsequently go out of their way to tell friend or family about this interviewer with extraordinary interpersonal skills. Usually include two or more elements of "positive speak" and expressions of nonverbal interest that are exceptionally warm.
4. Notably warm and makes effective connection via identifiable elements of both verbal and nonverbal connection
3. Clearly professional, respectful, and interested but minimal or ineffective specific verbal or nonverbal efforts to make a more personal connection.
2. For the most part professional and respectful. Absent of specific effective efforts at rapport building. Present are some comments, expressions, or nonverbal behaviors that might have a negative reception by at least some patients.
1. Absent are positive elements of relationship building. Present are clearly negative comments or expressions, which would leave most patients with negative feelings about the interviewer.

Agenda Setting—Global Criteria

5. Explores complete agenda at the beginning until the point that the patient says, "Nothing else." If several agenda items, prioritizes amongst them. Explores for additional agenda at end.
4. Explores complete agenda but may not summarize or prioritize or may not explore for more agenda at end.
3. Explores for agenda partially with at least two efforts at agenda setting. One can be at beginning and one at end.
2. Asks only once at the beginning, eg, "What brings you in today?" or "How can I be of help?" or at the end, "Is there anything else?"
1. Doesn't explore for agenda at beginning but begins addressing an established problem. Doesn't return to agenda at any point.

Information Management—Global Criteria

5. Begins interview with open-ended question and nondirected facilitation. Continues in this mode (with occasional closed-ended points of clarification) until most/all of patient's information about the condition has been expressed. Performs appropriate summary(s). Asks appropriate focused (closed) questions toward the end.
4. Begins with open-ended questions. Mixes open- and closed-ended questions. Uses some form of partial summary.
3. Uses some open-ended and closed-ended questions from the beginning. Doesn't summarize or does so weakly.
2. Mostly closed-ended questions. No summary or inadequate summary.
1. Mostly closed-ended questions. May use leading questions or repeats questions.

Active Listening to Understand the Patient's Perspective on Illness—Global Criteria

5. Very effective at identifying the patient's perspective on illness (PPI) (i e, what the patient thinks may be going on, the greatest concern about the problem, and the expectations for the visit). The PPI is repeatedly explored using active listening to understand the meaning behind the patient's "clues." Once the PPI is disclosed, these elements are acknowledged, normalized, and used as part of a plan to address the medical diagnosis and the PPI.
4. Demonstrates genuine interest in the PPI by using active listening at least part of the time. Does explore the clues initially but not always fully. Once identified, PPI will be partially addressed with some elements of acknowledgment, normalization, and building a plan based on the PPI.
3. Demonstrates some interest in the PPI through occasional exploration of clues (efforts may not be effective). May not pick up on clues but rather asks about the patient's ideas.
2. Fails to demonstrate effective interest in what the patient thinks may be going on, his/her greatest concern about the problem, and the expectations for the visit. (Clueless)
1. Actively discourages or devalues the PPI.

Addressing Feelings—Global Criteria

5. Responds to all opportunities to address feelings. When the patient expresses a feeling, these are acknowledged, normalized, or legitimized and are addressed with a follow-up, which at least explores how the patient would like these feelings to be addressed. Also seeks out the "potential feelings" when situations with high likelihood of feelings surface in the interview.
4. Acknowledges feeling when expressed and partially normalizes them and attempts a follow-up plan. Does not fully address potential feeling situations.
3. Acknowledges feelings but does not use the other skills mentioned above.
2. May superficially acknowledge one of a small portion of the feelings expressed. May not acknowledge any of the feelings of the case.
1. Comments or responds in a way that demeans, criticizes, or devalues patients' feelings.

Reaching Common Ground—Global Criteria

5. Works very effectively at bridging differences between the interviewer and the patient. Requires a full exploration of the PPI and use of the PPI to reach common ground. Uses a number of the more effective skills in reaching common ground, eg, full exploration of the PPI, decision analysis, reframing, patient-centered suggestions, criteria setting, brainstorming, compromise, etc. Avoids less effective methods, eg, use of authority, personal appeal, repetition of serious complications or chance of death. Would likely facilitate a desirable change in behavior toward health.
4. Demonstrates clear skills in reaching common ground. Does obtain most of the PPI and attempts to use at least some (but not all) of its elements in a plan. Uses a mix of strategies to reach the plan. Heavier use of the more effective skills, eg, full exploration of the PPI, decision analysis, reframing, patient-centered suggestions, criteria setting, brainstorming, compromise.
3. While does not connect the plan with PPI, uses a balanced mix of skills to reach common ground that includes at least one of the more effective strategies.
2. Does not use the patient's issues to help to resolve the difference. Uses more of the less effective strategies in trying to create a plan, eg, use of authority, personal appeal, repetition of serious complications or risk of death. For most patients this plan would not significantly affect the long-term behavior in question.
1. Uses less effective strategies almost exclusively. In missing the patient's issues and in using authority or threat, the patient would be unlikely to change long-term behavior and would probably leave upset with the interviewer's approach to problem solving.

Overall Global Rating Criteria

5. At the level of an experienced clinician who is expert in using all communications skills effectively. Skills demonstrated such that a patient would likely note such skills to friends and family.
4. Uses all communication skills effectively; minor suggestions for change are noted, which are unlikely to have measurable importance on encounter.
3. Uses most communication skills effectively; some interview behaviors present that, if modified, could lead to an even more effective impact on a real encounter.
2. Uses some communication skills effectively and others ineffectively; certain areas of communication might cause clinical problems (patient dissatisfaction.)
1. Inadequate communication skills; likely to create significant clinical problems (patient dissatisfaction)

In general, the numbers above translate into the following:5=exemplary, 4=very effective, 3=competent/adequate, 2=marginal, 1=needs improvement