

Use of the Kalamazoo Essential Elements Communication Checklist (Adapted) in an Institutional Interpersonal and Communication Skills Curriculum

BARBARA L. JOYCE, PHD
TIMOTHY STEENBERGH, PHD
ERIC SCHER, MD

Abstract

Background This study examined the psychometric properties of the Kalamazoo Essential Elements Communication Checklist (Adapted) (KEECC-A), which addresses 7 key elements of physician communication identified in the Kalamazoo Consensus Statement, in a sample of 135 residents in multiple specialties at a large urban medical center in 2008–2009. The KEECC-A was used by residents, standardized patients, and faculty as the assessment tool in a broader institutional curriculum initiative.

Methods Three separate KEECC-A scores (self-ratings, faculty ratings, and standardized patient ratings) were calculated for each resident to assess the internal consistency and factor structure of the checklist. In addition, we analyzed KEECC-A ratings by gender and US versus international medical graduates, and collected American Board of Internal Medicine Patient Satisfaction Questionnaire (PSQ) scores for a subsample of internal medicine residents ($n = 28$) to examine the relationship

between this measure and the KEECC-A ratings to provide evidence of convergent validity.

Results The KEECC-A ratings generated by faculty, standardized patients, and the residents themselves demonstrated a high degree of internal consistency. Factor analyses of the 3 different sets of KEECC-A ratings produced a consistent single-factor structure. We could not examine the relationship between KEECC-A and the PSQ because of substantial range restriction in PSQ scores. No differences were seen in the communication scores of men versus women. Faculty rated US graduates significantly higher than international medical graduates.

Conclusion Our study provides evidence for the reliability and validity of the KEECC-A as a measure of physician communication skills. The KEECC-A appears to be a psychometrically sound, user-friendly communication tool, linked to an expert consensus statement, that can be quickly and accurately completed by multiple raters across diverse specialties.

Introduction

Physician communication skills are associated with improved patient satisfaction, better health outcomes, greater adherence to treatment, and more active

self-management of chronic illnesses.^{1–7} The Accreditation Council for Graduate Medical Education, the American Board of Medical Specialties, and the Association of American Medical Colleges have underscored the importance of communication skills by including training and assessment in communication and interpersonal skills as one of the competency domains.^{8–10}

The Bayer-Fetzer Conference on Physician-Patient Communication in Medical Education convened authors of the major theoretical models of physician-patient communication and other important stakeholders to reach a consensus on the essential elements that characterized physician-patient communication. Their report resulting from this conference, termed the Kalamazoo Consensus Statement (Kalamazoo I), identified 7 key elements of communication in clinical encounters: build the relationship, open the discussion, gather information, understand the patient's perspective, share information, reach agreement, and provide closure.¹ It was hoped that by providing a common framework, this expert consensus

Barbara L. Joyce, PhD, is the Director of Instructional Design at Henry Ford Hospital and Clinical Associate Professor of Family Medicine at Wayne State University; **Timothy Steenbergh, PhD**, is Associate Professor of Psychology at Indiana Wesleyan University; and **Eric Scher, MD**, is Vice-President of Medical Education, Henry Ford Health System, Designated Institutional Official and Vice Chair of Internal Medicine, Henry Ford Hospital, and Clinical Associate Professor, Internal Medicine, Wayne State University.

The authors would like to thank Elizabeth Rider, MSW, MD, for her insight and work in this area. The authors would also like to thank Kathy Peterson for her help with the data management on this project. This project was funded by Henry Ford Hospital Department of Medical Education; no outside funding was received.

Corresponding author: Barbara L. Joyce, PhD, Henry Ford Hospital, Department of Medical Education, CFP 046, Detroit, MI 48202, 313.916.9161, Bjoyce2@hfhs.org

Received February 9, 2010; revision received March 22, 2010; accepted April 13, 2010.

DOI: 10.4300/JGME-D-10-00024.1

statement would facilitate the development of communications curriculum and assessment tools in medical education. Later, The Kalamazoo II report recommended specific assessment methods to evaluate communication skills, including direct observation with real patients, ratings of simulated encounters with standardized patients (SPs), ratings of video or audiotaped interactions, patient surveys, and knowledge/skills/attitudes examinations.¹¹

The original Kalamazoo Essential Elements Checklist included 23 items assessing subcompetencies identified in the Kalamazoo I report. However, its scaling options (done well, needs improvement, not done) limited a rater's ability to distinguish among the range of physician communication skills. Furthermore, it required considerable time for rating. In response to these limitations, Calhoun et al¹² adapted the Kalamazoo Essential Elements Checklist by replacing the original response options with a 5-point Likert scale that allowed raters to evaluate each communication skill on a continuum from "poor" to "excellent," and shifted to global ratings for the 7 essential elements. This tool will be subsequently referred to as the KEECC-A.

Initial studies of the KEECC-A suggest that it is a flexible tool with psychometric data to support its use in some medical education settings. Rider et al² studied fellows during a simulated family meeting using the KEECC-A and a gap analysis as part of a multisource assessment of communication skills. Investigators in that study added 2 additional dimensions (demonstrates empathy and communicates information accurately) to the KEECC-A instrument and found a Cronbach α value of .84 for the original 7 dimensions and a Cronbach α value of .87 for the 9 dimensions of their modified tool. These data provided evidence for the internal consistency of the measure when completed by these peer raters and suggested that one strength of the KEECC-A is ease of use by multiple raters.¹²

The usability of the KEECC-A has also been highlighted. Schirmer et al¹³ reported that the KEECC-A helped less experienced faculty raters focus on 7 key elements of communication. Furthermore, Calhoun et al¹² found the average time to complete the KEECC-A, plus an additional 2 dimensions, was 7 (± 2.7) minutes, suggesting that this tool is feasible for faculty to use.

We developed an institutional curriculum in communication skills for first-year core residency programs using the Kalamazoo framework to guide curricular development and assessment. This curriculum focused on 3 key topic areas: informed consent, disclosure of errors, and sharing bad news. Each of these 3 communication topic areas contained an online module, a small group discussion facilitated by the program director or key clinical faculty, and a 3-station objective structured clinical examination (OSCE) for each topic area (9 OSCEs total). Residents, faculty, and SPs used the KEECC-A as the assessment tool for the OSCEs.

This article extends and expands on the work done by Rider et al¹⁴ and Calhoun et al¹² by looking at evidence for the reliability and validity of the KEECC-A using ratings from SPs, faculty, and resident self-ratings. In addition, scores on the KEECC-A and the American Board of Internal Medicine Patient Satisfaction Questionnaire (PSQ) were analyzed for a subgroup of residents in order to examine evidence for the measure's convergent validity. The KEECC-A has the potential to fill the need for a communication tool that is linked to an expert conceptual framework, is brief, and can be used for assessment of communication skills by a wide variety of raters and specialties.

Methods

Participant Characteristics

Participants (N = 135) included first-year and second-year residents (who were new to our institution) from 16 residency programs in a large teaching hospital during the academic year 2008–2009. Fifty-nine (43.7%) were primary care residents, 53 (39.3%) were hospital-based residents, and 23 (17.0%) were surgery residents. Women (n = 50) made up 37% of the sample. Fifty-nine percent of residents were graduates of US medical schools. All participants gave informed consent and the Henry Ford Health System Institutional Review Board approved the study.

Instruments

The KEECC-A is a 7-item rating scale with each item corresponding to 1 of the 7 essential elements of physician communication that were identified by an expert consensus panel.¹ Ratings are made on a 5-point Likert scale (1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent). Responses to the 7 items are summed to provide a total communication score, with higher scores representing greater communication skill. The KEECC-A can be completed by an observer (eg, attending physician, SP) or self-rated by the physician.

The American Board of Internal Medicine PSQ is a 10-item measure of patients' perceptions of their physician's communication and interpersonal skills. Previous work has provided reliability and validity evidence for the PSQ in a resident sample.¹⁵ For the purposes of this study, we focused only on the physician communication skills factor as identified by Yudkowsky et al.¹⁵ Mean item scores were computed for the communication factor so that scores could range from 1 to 5, with higher scores indicating greater communication skill.

Procedure

Evidence for the reliability and validity of the KEECC-A was drawn from residents' OSCE data from the first module to be implemented. Three different KEECC-A ratings were generated for each participant in this module. A faculty member, an SP, and the participating resident rated the

resident's performance using the KEECC-A. Residents self-assessed their performance on the OSCE immediately afterward using the KEECC-A. Standardized patients (N = 15), who had formalized training in using the KEECC-A, assessed the residents immediately afterward using the KEECC-A and then provided verbal feedback. Faculty (N = 25) assessed residents later by reviewing the videotape of their OSCE and providing written and verbal feedback to the resident in a mentoring meeting. Faculty were given clear instructions in a faculty debrief guide about rating the residents on the KEECC-A prior to reviewing the self-assessment or SP ratings and faculty were also given key learning points to discuss.

Three separate KEECC-A scores (self-ratings, faculty ratings, and SP ratings) were calculated for each resident by adding the ratings of the 7 items for each respective rater using the measure. Each participant had a self-rating score, an SP-rated score, and a faculty-rated score on the KEECC-A. Scores could range from 7 to 35.

The PSQ scores were gathered for a subsample of internal medicine residents (n = 28) to examine the relationship between this measure of communication skills and the KEECC-A ratings. Patients completed the instruments following their visit with resident physicians, and the mean rating for resident physicians across patients was used to determine PSQ communication scores.

Results

KEECC-A Scores Across the 3 Rating Groups

Residents' cumulative self-ratings on the KEECC-A ranged from 14 to 35 with a mean (SD) of 26.87 (5.01). Most (79.8%) residents rated their communication skills as "good" (21) or better. Faculty ratings ranged from 13 to 35 with a mean of 25.25 (5.08). A majority of faculty (79.9%) rated residents' skills as "good" or higher, and 27.9% of faculty provided "very good" or "excellent" ratings. Cumulative SP ratings ranged from 13 to 35 with a mean of 21.72 (4.53). A smaller proportion of SPs (54.1%) rated residents as "good" or better, and only 15.6% provided ratings of "very good" or "excellent."

We examined the relationship between faculty ratings, SP ratings, and self-ratings. The strongest correlation was between faculty ratings and SP ratings ($r_{132} = .31, P < .001$), providing some evidence for interrater reliability. The KEECC-A self-ratings were not significantly related to faculty ratings ($r_{122} = .09, P > .05$) or SP ratings ($r_{123} = .12, P > .05$). These low correlation coefficients may be partly due to restricted range in faculty ratings and self-ratings, which tended to be uniformly high. Only 3.9% of residents rated themselves below an average rating of 3 (good), and nearly half (48.2%) of residents rated themselves at 4 (very good) or above.

To further explore whether residents accurately evaluated their communication skills, we examined their

TABLE 1
TWO-WAY ANALYSES OF VARIANCE FOR KALAMAZOO ESSENTIAL ELEMENTS COMMUNICATION CHECKLIST-ADAPTED RATINGS ACROSS GENDER AND UNITED STATES MEDICAL GRADUATE (USMG) VERSUS INTERNATIONAL MEDICAL GRADUATE (IMG) GROUPS

| Score | Source | df | F | P Value |
|-----------------|--------------------------|-----|------|---------|
| Faculty | Gender | 1 | 1.94 | .17 |
| | US/IMG ^a | 1 | 7.75 | .01 |
| | Interaction ^b | 1 | .97 | .33 |
| | Error | 133 | | |
| SP ³ | Gender | 1 | .17 | .68 |
| | US/IMG ^a | 1 | 1.91 | .17 |
| | Interaction ^b | 1 | 3.11 | .08 |
| | Error | 131 | | |
| Self-rating | Gender | 1 | 1.91 | .17 |
| | US/IMG ^a | 1 | .88 | .35 |
| | Interaction ^b | 1 | .35 | .55 |
| | Error | 121 | | |

Abbreviations: df, degrees of freedom; SP, standardized patient.

^a US/IMG represents the comparison between US and international medical graduates.

^b Interaction denotes the interaction effect of Gender × US/IMG.

self-ratings on the KEECC-A using SP ratings as a standard. The mean self-ratings of residents with the weakest communication skills (bottom quartile on the SP-rated KEECC-A, mean = 26.81 [5.74]) were not significantly different than self-ratings of residents with the strongest communication skills (top quartile, mean = 27.41 [4.95]), $t_{61} = .45, P > .05$. These data suggested that there was no difference in how residents rated themselves, despite the fact that SPs clearly saw a difference in these 2 groups. Faculty ratings showed greater correspondence with SP ratings. Faculty-rated KEECC-A scores were significantly higher for residents in the top quartile of the SP-rated KEECC-A (mean = 27.74 [5.26]) relative to residents, whose SP ratings were in the bottom quartile (mean = 23.92 [4.44]), $t_{66} = 3.91, P < .01$.

Comparisons of Demographic Groups

We used 3 separate 2-way analyses of variance to test for differences in KEECC-A ratings (self-ratings, SP ratings, and faculty ratings) between genders and between international medical graduates (IMGs) versus US medical graduates (USMGs). The analyses of variance results are presented in TABLE 1 and reveal no significant main effects for gender across the 3 different KEECC-A ratings. Main effects for IMGs versus USMGs were only significant for faculty

TABLE 2
FACTOR LOADINGS FOR FACULTY, STANDARDIZED PATIENT (SP), AND SELF-RATED KALAMAZOO ESSENTIAL ELEMENTS COMMUNICATION CHECKLIST-ADAPTED (KEECC-A) ITEMS

| KEECC-A Item | Faculty-Rated KEECC-A | SP-Rated KEECC-A | Self-Rated KEECC-A |
|---------------------------------------|-----------------------|------------------|--------------------|
| Builds relationships | .70 | .84 | .84 |
| Opens the discussion | .83 | .84 | .82 |
| Gathers information | .84 | .86 | .86 |
| Understands the patient's perspective | .84 | .81 | .86 |
| Shares information | .77 | .70 | .86 |
| Reaches agreement | .71 | .77 | .87 |
| Provides closure | .71 | .72 | .89 |

ratings, with faculty rating USMGs (mean = 26.46 [5.34]) significantly higher than IMGs (mean = 23.46 [4.11]).

Although there were no statistically significant interaction effects on SP ratings, there was a nonsignificant trend for SPs to rate male IMGs significantly lower than male USMGs ($P = .08$).

Internal Reliability

Cronbach α values were .89 for faculty ratings, .90 for SP ratings, and .94 for self-ratings, suggesting a high degree of internal consistency across KEECC-A items in this sample. Each of the individual items for each version of the scale was found to be strongly related to the overall measure.

Factor Analysis

Three principal components analyses were conducted to determine the factor structure of the 3 groups of KEECC-A ratings. Analysis of the faculty-rated KEECC-A revealed a clear single-factor solution based on the scree test and second eigenvalue below 1.0. All items loaded at .70 or higher on the factor. Loadings for each item are presented in TABLE 2. The factor explained 59.86% of the variance in the faculty-rated KEECC-A items.

Similar scree tests and eigenvalues were found for SP- and self-rated KEECC-A items, supporting single-factor solutions for both of these versions of the measure. Item loadings for the SP-rated KEECC-A were similar to those for the faculty-rated version; these are presented in TABLE 2. Single-factor models explained 62.47% of the variance in SP-rated KEECC-A items and 73.34% of the variance in self-ratings.

Relationship Between PSQ Scores and KEECC-A Ratings

Mean PSQ communication scores were uniformly high, 4.54 (.34). Pearson correlation coefficients between PSQ scores and the 3 different KEECC-A ratings revealed no significant relationship, all $P > .05$. However, given that patients completing the PSQ consistently scored residents very high, the restricted range of the PSQ scores limited our

ability to discover a relationship between the PSQ and the KEECC-A.

Discussion

In our sample of 135 residents from multiple specialties, the KEECC-A ratings of faculty, SPs, and residents demonstrated high internal consistency ranging from $\alpha = .89$ for faculty ratings to $\alpha = .94$ for residents' self-ratings. These reliability coefficients are noteworthy, given the small number of KEECC-A items. Furthermore, they are consistent with or higher than those previously reported on the KEECC-A (eg, $\alpha = .84$ reported by Calhoun et al¹² in 2009, and $\alpha = .88$ for the original KEECC reported by Schirmer et al¹³ in 2005).

Interrater reliability scores across the 3 ratings for each resident on the KEECC-A revealed a mixed pattern. Ratings by SPs and faculty were significantly, although modestly correlated, providing some evidence for interrater reliability. Residents' self-ratings did not correlate with either faculty ratings or SP ratings, and were significantly higher than SP or faculty ratings. This may suggest a weakness in residents' self-evaluation skills, rather than a weakness in the KEECC-A, and a finding that health professionals have difficulty accurately assessing their performance is consistent with previous work in this area.¹⁶⁻¹⁸

Faculty rated USMGs communication skills significantly higher than those of IMGs; however, there were no differences between groups across either self-ratings or SP ratings. We are unclear about what these data suggest, given the inconsistency between faculty ratings and corresponding ratings by SPs and residents themselves. Further research in this area is recommended.

Factor analyses of the 3 sets of KEECC-A ratings produced a consistent 1-dimensional factor structure suggesting that, in this sample, items were highly related to a general communication factor. These results, which provide the first glimpse at the factor structure of the KEECC-A, offer evidence for the construct validity of the

measure. Furthermore, content validity has been established by the very nature of the items, which were developed by experts in the area whose recommendations were based on empirical studies examining communication.¹ Future studies examining the factor structure of this measure may be helpful in clarifying whether the unidimensional structure that we found is unique to our sample.

One of the limitations of this study was its failure to find convergent validity evidence. Unfortunately, without a “gold standard” for physician communication skills, establishing convergent validity for measures such as the KEECC-A remains a challenge.¹¹ Our attempt to examine the relationship between KEECC-A and another validated communication skills tool, the PSQ, was thwarted because of the substantial range restriction in PSQ scores.

Another potential limitation of the study was that all participating residents received a communication skills training module, which likely influenced self-assessment scores (as we would expect). Residents may have perceived that they were doing a better job of communicating because they had just undergone training and were sensitized to the demand characteristics inherent in educational processes such as these. Likewise, the elevated scores on physician ratings may also represent a potential rating bias associated with evaluating the effectiveness of one’s own student.

Conclusions

This study provides reliability and validity evidence for the KEECC-A as a measure of physician communication skills in a sample of residents from multiple specialties in a large urban medical center. The KEECC-A provides a user-friendly communication tool, linked to an expert consensus statement, that can be quickly and accurately completed by a variety of raters. Combining the conceptual framework outlined by the Kalamazoo Consensus Statement with the KEECC-A allows for the robust development of communication curriculum and assessment in graduate medical education. Further research to examine evidence of convergent or predictive validity of the KEECC-A is needed.

References

- 1 Makoul G. Essential elements of communication in medical encounters: the Kalamazoo consensus statement. *Acad Med*. 2001;76(4):390–393.
- 2 Rider EA, Nawotniak RH, Smith G. *A Practical Guide to Teaching and Assessing the ACGME Core Competencies*. Marblehead, MA: HCPro, Inc; 2007:1–85.
- 3 Chen JY, Tao ML, Tisnado D, et al. Impact of physician-patient discussions on patient satisfaction. *Med Care*. 2008;46(11):1157–1162.
- 4 Heisler M, Bouknight RR, Hayward RA, Smith DM, Kerr EA. The relative importance of physician communication, participatory decision making, and patient understanding in diabetes self-management. *J Gen Intern Med*. 2002;17(4):243–252.
- 5 Heisler M, Tierney E, Ackermann RT, et al. Physicians’ participatory decision-making and quality of diabetes care processes and outcomes: results from the triad study. *Chronic Illn*. 2009;5(3):165–176.
- 6 Makoul G, Curry RH. The value of assessing and addressing communication skills. *JAMA*. 2007;298(9):1057–1059.
- 7 Schoenthaler A, Chaplin WF, Allegrante JP, et al. Provider communication effects medication adherence in hypertensive African Americans. *Patient Educ Couns*. 2009;75(2):185–191.
- 8 Accreditation Council for Graduate Medical Education. Common program requirements: general competencies. Available at: <http://www.acgme.org/outcome/comp/compCPRL.asp>. Accessed September 23, 2009.
- 9 American Board of Medical Specialties. Maintenance of certification competencies and criteria. Available at: http://www.abms.org/Maintenance_of_Certification/MOC_competencies.aspx. Accessed September 23 2009.
- 10 Association of American Medical Colleges. Learning objectives for medical school education: guidelines for medical schools. Available at: https://services.aamc.org/publications/showfile.cfm?file=version87.pdf&prd_id=198&prv_id=239&pdf_id=87. Accessed September 23, 2009.
- 11 Duffy FD, Gordon GH, Whelan G, et al. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med*. 2004;79(6):495–507.
- 12 Calhoun AW, Rider EA, Meyer EC, Lamiani G, Truog RD. Assessment of communication skills and self-appraisal in the simulated environment: feasibility of multirater feedback with gap analysis. *Simul Healthc*. 2009;4(1):22–29.
- 13 Schirmer JM, Mauksch L, Lang F, et al. Assessing communication competence: a review of current tools. *Fam Med*. 2005;37(3):184–192.
- 14 Rider EA, Hinrichs MM, Lown BA. A model for communication skills assessment across the undergraduate curriculum. *Med Teach*. 2006;28(5):e127–134. <http://informahealthcare.com/doi/abs/10.1080/01421590600726540>. Accessed October 10, 2009.
- 15 Yudkowsky R, Downing SM, Sandlow LJ. Developing an institution-based assessment of resident communication and interpersonal skills. *Acad Med*. 2006;81(12):1115–1122.
- 16 Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA*. 2006;296(9):1094–1102.
- 17 Regehr G, Eva K. Self-assessment, self-direction, and the self-regulating professional. *Clin Orthop Relat Res*. 2006;449:34–38.
- 18 Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. *Adv Health Sci Educ Theory Pract*. 2002;7(1):63–80.