Validity Evidence Supporting the Comunicación y Habilidades Interpersonales (CAI) Scale for Medical Spanish Communication and Interpersonal Skills Assessment

Pilar Ortega, MD, Nathaniel R. Moxon, Anuj K. Chokshi, Cristina Pérez-Cordón, PhD, and Yoon Soo Park, PhD

Abstract

Purpose
While many schools teach medical Spanish to improve communication with the growing Spanish-speaking population, a standardized assessment methodology is lacking. This study evaluates validity evidence for a newly developed Spanish Communication and Interpersonal Skills (CIS) scale.

Method
The authors developed the Comunicación y Habilidades Interpersonales (CAI, pronounced /kəˈmuniˌkɑːsən i ˈhæbɪldəs ɪntərˈpɜːrsənləs/) scale by culturally and linguistically adapting a 14-item validated English CIS scale. Sources of validity evidence were gathered targeting content, response process, internal structure, relations to other variables, and consequences. The CAI was administered as part of an objective structured clinical examination at 2 medical schools from April to June 2020.

Results
The authors mapped CAI items onto USMLE Step 2 CIS behavioral functions and medical Spanish learner competencies to examine content validity. The mean item difficulty and item discrimination estimates are 2.86 (SD = 0.34) and 0.81 (SD = 0.08), respectively, demonstrating good psychometric properties at the item level. Internal-consistency reliability for a single case was 0.96. Learner variance accounted for 12% of total variance, demonstrating ability to differentiate higher and lower performing learners; the person–case interaction accounted for 44% of total variance, indicating strong case specificity.

Conclusions
The CAI has higher case specificity than previously reported in English, suggesting that the nuanced contextual elements of individual cases may matter even more when learners are using limited language skills. Implementing the CAI more widely may promote inclusion of patient-centered communication skills as part of curricular content and activities. This study’s validity evidence may inform the development of a future medical Spanish certification examination.

Communication and interpersonal skills (CIS) represent required elements of medical student core competencies and are critical for successful transition to residency and medical practice. The Liaison Committee on Medical Education requires that accredited medical schools teach cultural competence, health care disparities, and communication skills with patients and their families.1 The Association of American Medical Colleges integrates CIS throughout its published Core Entrustable Professional Activities for Entering Residency.2 Additionally, the Accreditation Council for Graduate Medical Education requires teaching residents effective communication “across a broad range of socioeconomic and cultural backgrounds” to assess goals and determine care.3

The stated requirements for communication skills with diverse populations are at odds with the reality that United States medical schools and accrediting bodies only routinely teach and assess CIS in English. For example, the United States Medical Licensing Examination (USMLE) Step 2 Clinical Skills (CS) assesses candidate CIS, but only with English-speaking patients.4 Given that over 350 languages are spoken by the U.S. general population,5 the lack of CIS assessment in non-English languages may result in unchecked use of limited skills and inadvertently contribute to structural barriers to health care for populations with limited English proficiency (LEP).6

Spanish is the largest and fastest growing non-English language in the U.S., with an estimated 40 million individuals speaking Spanish at home, of which nearly 40% report LEP.7 Hispanic/Latinx individuals are more likely to die from preventable causes including diabetes, obesity, and liver disease.8 Relatedly, persons with LEP experience better outcomes if treated by a language-concordant physician (a physician who speaks their language).9 It is anticipated that outcomes for Spanish-speaking patients would improve if consistent standards for medical Spanish education and assessment were implemented with robust validity evidence.10

While student demand has resulted in a growing number of medical Spanish programs at U.S. medical schools, few programs report a method of skills assessment.11,12 According to the most recent data from a 2019 national survey in which 125 U.S. medical schools participated, 78% of schools reported some medical Spanish education, yet only 57% reported conducting learner postcourse skills assessment.13 Of schools whose medical Spanish education is considered extracurricular (as opposed to curricular elective or required courses), the use of assessment is even less frequent (20%).12 As a result, medical students often use unassessed language skills with
patients out of perceived necessity, risking patient safety due to potential miscommunication. A recent push for standardization has resulted in the identification of medical Spanish core competencies, performance objectives, and domains that should be evaluated by an assessment examination. The domains include general oral proficiency, listening comprehension, oral medical proficiency, and CIS. However, the lack of a unified medical Spanish assessment examination and validated rater tools remains an unresolved problem.

Similar to USMLE Step 2 CS in English, objective structured clinical examinations (OSCEs) using standardized patients (SPs) can be used to assess multiple domains in medical Spanish communication. Importantly, the CIS domain should include patients’ perspectives. Several rating instruments for the assessment of English CIS have been developed, revised, and shown to have validity evidence. However, a CIS scale for use in medical Spanish learner assessment with validity evidence has never been published.

While some Spanish-speaking SPs may have English skills, they may not be as comfortable completing rating forms in English as they would be in Spanish. Moreover, CIS includes culturally nuanced abilities such as building rapport, use of verbal and nonverbal cues, and demonstration of respect/empathy, which have important differences in the Hispanic/Latinx culture and Spanish language. Thus, there is a need for a CIS scale for medical Spanish learners that reflects patient–physician communication elements in a linguistically and culturally appropriate manner.

This multisite study evaluates validity evidence for a newly developed CIS scale at 2 medical schools for use by Spanish-speaking patients in the assessment of medical student clinical skills in Spanish. We use Messick’s unified validity framework to gather and examine validity evidence relating to content, response process, internal structure, relations to other variables, and consequences following Messick’s unified validity framework.

Method

We developed a new assessment, called Comunicación y Habilidades Interpersonales (CAI) (pronounced /ki/-/hey-bell-ee-daw-dews/- in all its standard geographic variants). This facilitates pronounceability and functionality for use in U.S. medical schools in which the primary language of medical education is English.

Iterative revisions of the tool took place between October 2019 and April 2020. We then administered the CAI and collected data from April to June 2020 at Northwestern University (NU) Feinberg School of Medicine and University of Illinois at Chicago (UIC) College of Medicine as part of OSCEs administered during a medical Spanish course. We obtained data for validity evidence of the CAI from the following sources: (1) feedback from patient and expert reviewers for iterative tool revisions, (2) CAI scores from medical students’ Spanish SP encounters, (3) student performance on other medical Spanish course metrics, and (4) feedback from students evaluated using the CAI.

Data compilation and analyses were conducted using Stata 16 (Stata Corp, College Station, Texas). This study was determined to be exempt by the institutional review boards of UIC (Protocol # 2019-0945) on September 13, 2019, and NU (Protocol # STU00212479) on November 17, 2020.

We gathered sources of validity evidence for CAI targeting content, response process, internal structure, relations to other variables, and consequences following Messick’s unified validity framework.

Content

We conducted a review of literature to seek potential tools for CIS evaluation of U.S. medical students in Spanish. To be useful for U.S. medical Spanish, the tool’s content should align well with target medical Spanish learner competencies as well as existing English CIS behavioral functions in U.S. medical education. To achieve these content goals, we identified the Revised University of Illinois Communication and Interpersonal Skills (RUCIS) scale as a tool which has previously demonstrated validity evidence and sound psychometric results. Using RUCIS as a frame of reference, we culturally and linguistically adapted its content to align with medical Spanish course objectives for use by Spanish-speaking U.S. patients or SPs.

A linguist (C.P.-C.) first translated and adapted the RUCIS into Spanish, and a physician/medical Spanish educator (P.O.) reviewed the initial translation for alignment with the target clinical skills competencies of medical Spanish courses. A certified Spanish-English translator then reviewed and suggested edits to the revised scale. We made additional edits based on feedback from 2 monolingual Spanish speakers and 2 experienced, native Spanish-speaking SPs from different Hispanic/Latinx nationalities. We asked the 2 monolingual Spanish speakers to imagine using the scale to evaluate their own physicians following a medical encounter. Inclusion of diverse nationalities was important because national origin can play a role in various factors, which can affect medical communication, such as pronunciation, prosody, vocabulary, and culture.

We developed 2 medical Spanish cases (abdominal pain and shortness of breath) for which students were assessed using the CAI. The cases were first designed as a formative and summative assessment of student medical Spanish performance and reported in 2017 by Ortega et al. We subsequently refined the cases to include more complex interpersonal elements that would enable a more nuanced understanding of learner CIS. For example, we added elements to assess skills pertaining to providing clear explanation, sensitive subject matters, and discussion of options/plans.

Response process

To study the validity evidence of the CAI based on response processes, we selected qualified and experienced SPs, trained them to use the CAI, and evaluated SP feedback on the tool. To ensure the authenticity of the encounter, SPs who are recruited for medical Spanish cases must report a language preference of Spanish
during medical encounters. Two such experienced native Spanish-speaking SPs were trained to use the CAI during a 2-hour training session in addition to their prior SP training. To examine trends in data, including assessor use of rating scale in CAI, we examined descriptive statistics at the item, case, and overall assessment levels.

Internal structure
We examined case-level internal-consistency reliability using Cronbach's alpha and also conducted item analysis (item difficulty and item discrimination) to gather psychometric data for the CAI. In addition, we conducted generalizability theory (G-theory) analysis to examine variance components using a fully crossed design, person x case x items; following the same G-theory design used in prior RUCIS studies. Projections in reliability were estimated using Decision studies.

Relation to other variables
We assessed the relationship to other variables by correlating students' scores on the CAI with their scores on other aspects of the Spanish encounter, including the case note score by the medical Spanish faculty and the SP checklist score, and other performance metrics of the medical Spanish course, including their score on a knowledge assessment exam and a self-reported confidence assessment.

Consequences
We compared pass–fail rates using Angoff cut scores, previously determined for RUCIS. After students received their scores and performance feedback for the SP encounter, including CAI scores, they each completed a guided reflection written assignment. The reflection assignment asked specific questions about their self-perceived strengths and limitations, their opinion about utility of the overall feedback they received following the encounter, and the specific utility of the feedback provided by their CAI scale results. The reflection form also provided a free-text area for any other learner feedback or comments. Two members of the research team then examined qualitative feedback and reflections from students by compiling all comments that were related to the CAI scale itself, self-perception of interpersonal skills development, or their performance and identified common themes.

Results
We collected data for a total of 34 medical students enrolled in a medical Spanish course at 2 medical schools (NU: n = 17; UIC: n = 17). All students participated in 2 Spanish SP encounters which were used for formative and summative assessment. Twenty-nine (86%) participants were fourth-year medical (M4) students, 4 (12%) were third-year medical (M3) students, and 1 (3%) was a physician who had just graduated from medical school. Forty-seven percent of participants (16) self-identified as female and 53% (18) as male. Average participant age was 25.8 years (SD = 2.27, range 23–33 years). Ethnic self-identification of participating students showed that 27% (9) identified as Hispanic/Latinx, significantly higher than the current national average for U.S. medical students (10%), and included national origins from Cuba, Colombia, Dominican Republic, Ecuador, El Salvador, Mexico, and Puerto Rico.

All students were required to have a general Spanish level in the low–intermediate range or above to enroll in the medical Spanish course, as recommended by expert consensus. Students varied in the extent of their prior Spanish experience. The majority of the Hispanic/Latinx participants (8 of 9 [88%] Hispanic/Latinx students, or 24% of 34 total students) reported being heritage Spanish speakers, meaning that Spanish was spoken in their household in some capacity while growing up. Of 34 participants, 2 students (6%) reported participating in prior medical Spanish education, 8 (24%) reported advanced Spanish education in college, 10 (29%) reported basic college Spanish coursework, and 25 (74%) reported participation in high school Spanish.

Content
We made iterative revisions to the CAI for ease of use in Spanish by Hispanic/Latinx patients. Reviewers' nationalities included Mexico, Nicaragua, and Spain. The translator was selected for her experience in multiple varieties of Spanish including regionalisms and sentence structure, to ensure that the final scale was appropriately accessible and inclusive of a wide diversity of Spanish speakers. Several CAI items represent departures from the RUCIS scale due to cultural adaptations and linguistic changes. For example, we added more extensive general instructions to guide patients on how to use the tool. In Spanish, some concepts are expressed using varied regional words or phrases, so, when appropriate, we included synonyms for key behaviors to maximize clarity and inclusiveness. For instance, under the friendly communication item, the English word "rudely" has no definitive literal translations but rather can be described using many Spanish words; the team opted for a combination of 2 terms (grosero y maleducado). As expected due to the wordier nature of the Spanish language, the CAI is lengthier than the English RUCIS. Table 1 describes some of the key cultural and linguistic adaptations for each CAI item's development. The final CAI scale is available as Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/B153.

We mapped CAI items into the 5 USMLE Step 2 CIS behavioral functions as well as the 5 medical Spanish learner competencies and associated learner performance objectives to examine content validity evidence. Table 2 provides a list of items and anchors. Classifications were determined by consensus of a physician/medical Spanish educator (P.O.), a linguist/language assessment expert (C.P.-C.), and a psychometrician (Y.S.P.).

Response process
SPs who participated in the study were 1 male and 1 female native Spanish speakers of Mexican origin with multiple years of experience as trained medical Spanish SPs. The SPs reported that the CAI was straightforward and easy to use. One said that it was a "good way to express what made me feel more or less comfortable with the doctor." Another SP stated that sometimes students "are very good at asking the medical questions ... but might not understand why I felt withdrawn or intimidated by how they spoke to me. It's nice to have a form that captures where they can improve." During their use of the CAI, SPs had the opportunity to provide free-text comments to explain their responses and suggest edits or changes to the scale, which allowed further scale refinement.

Across the CAI, assessors used the full range of the 4-point rating scale for all items, except items 1 (friendly
Table 1
Summary of Cultural and Linguistic Modifications During Development and Iterative Revisions of the CAI Scale

<table>
<thead>
<tr>
<th>CIS item</th>
<th>CAI cultural/linguistic adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructions</td>
<td>We added more extensive explanation to help guide patients in answering the questions properly.</td>
</tr>
<tr>
<td>1. Friendly communication</td>
<td>Concept of rude behavior can be communicated in multiple ways in Spanish, so we added 2 descriptors. We elaborated on the Spanish description of “warm and friendly environment” for clarity.</td>
</tr>
<tr>
<td>2. Respectful treatment</td>
<td>Being treated “as an inferior” can be communicated in multiple ways in Spanish, so 2 descriptors were added.</td>
</tr>
<tr>
<td>3. Listening to my story</td>
<td>Given that some medical Spanish learners have limitations in their listening comprehension of the patient, it was particularly important to integrate language regarding whether the patient felt that the clinician understood him/her and confirmed important elements of the patient’s story.</td>
</tr>
<tr>
<td>4. Honest communication</td>
<td>As a way to communicate respect and empathy. For example, the words “negative” and “positive” can have conflicting and confusing meanings in health care (a positive test result vs a good test result), and our reviewers felt explaining this in Spanish using other words would be best.</td>
</tr>
<tr>
<td>5. Interest in me as a person</td>
<td>Given the particular importance of family and familial relationships for Spanish-speaking patients, we adapted answer choices to include interest in the patient’s family and their well-being.</td>
</tr>
<tr>
<td>6. Discussion of options/plans</td>
<td>Given varied medical decision-making styles among Spanish speakers (some patients prefer to be more or less involved in medical decision making and may rely on doctors or family for support), we adapted answer choices to include appropriate response to the patient’s desired involvement in making decisions.</td>
</tr>
<tr>
<td>7. Encouraging my questions</td>
<td>We elaborated on the descriptors of encouraging questions which require more extensive explanations for clarity in Spanish. The term “question” can be stated in several ways in Spanish, so we included several terms and descriptors.</td>
</tr>
<tr>
<td>8. Providing clear explanation</td>
<td>We elaborated on the descriptors of providing clear explanations in Spanish. Given that some medical Spanish learners have limitations in their listening comprehension of the patient, it was particularly important to integrate language regarding whether the patient felt that the clinician’s explanations were in an understandable register.</td>
</tr>
<tr>
<td>9. Physical exam</td>
<td>The term “physical examination” can be communicated in multiple ways in Spanish, so we included 2 descriptors. We avoided terminology around positive or negative findings to avoid confusion and instead provided a more descriptive explanation of physical exam findings.</td>
</tr>
<tr>
<td>10. Appropriate vocabulary</td>
<td>We elaborated on the descriptors of appropriate vocabulary in Spanish.</td>
</tr>
<tr>
<td>11. Sensitive subject matters</td>
<td>The term “sensitive” can be communicated in multiple ways in Spanish, so we included several descriptors. We expanded upon possible examples of showing empathy that reflect common expressions in Spanish.</td>
</tr>
<tr>
<td>12. Concluding the encounter</td>
<td>We elaborated on the descriptors of worsening symptoms and return precautions. We provided examples of structural barriers to care that clinicians should be cognizant of such as inaccessibility to pharmacy or medication costs.</td>
</tr>
<tr>
<td>13. Receptiveness to feedback</td>
<td>We elaborated on the descriptors of how feedback can be given and received.</td>
</tr>
<tr>
<td>14. Do I want to see you again as my personal physician?</td>
<td>We elaborated on the descriptors of how pleased the patient may be with the physician. Given the particular importance of family and familial relationships for Spanish-speaking patients, we added whether the patient would recommend this physician to family members (in addition to recommending to friends).</td>
</tr>
</tbody>
</table>

Abbreviations: CAI, Comunicación y Habilidades Interpersonales (Spanish for “Communication and Interpersonal Skills”); CIS, Communication and Interpersonal Skills.

Interpersonal Communication Skills

Internal structure
CAI item difficulty and item discrimination results are shown by case in Table 3. Mean item difficulty and item discrimination estimates are 2.86 (SD = 0.34) and .81 (SD = 0.08), respectively, demonstrating good psychometric indices at the item level. The case-level Cronbach’s alpha reliabilities for shortness of breath and abdominal pain are 0.96 and 0.97, respectively, showing excellent consistency in case-level CAI scores.

Using generalizability theory, we examined facets contributing to variability in scores. Learner variance accounted for 12% of total variance, demonstrating that the CAI assessment is able to differentiate high and low performing learners: the person–case interaction accounted for 44% of total variance, indicating evidence of strong case specificity (learner performance varies by case administered). Item variance accounted for 15% of total variance.

The combined Φ-coefficient reliability across 2 cases using G-theory was 0.33, consistent with prior CIS reliability results for 2 cases; the corresponding G-coefficient reliability for 2 cases was 0.34. Projections in reliability using Decision study indicate that up to 12 cases should be administered to achieve Φ-coefficient reliability > 0.70.

Relations to other variables
Table 4 displays the relation between students’ CAI scores and their scores on other aspects of the medical encounters as well as other course performance metrics. Specifically, students’ overall CAI scores (sum score from both SP cases) were inversely correlated with their performance on the SP checklist questionnaire (P < .05). This means that students who received lower scores on SP checklist items, which assess factual information gathering (e.g., student asked about location of my communication) and 2 (fostering relationship), which used categories 2 to 4. Overall CAI performance was 72% (SD = 13%). There were no significant differences in performance by case, P = .793. Also, there were no significant differences by school, reported race/ethnicity, or gender, all P > .05.
<table>
<thead>
<tr>
<th>CAI item</th>
<th>RUCIS item</th>
<th>USMLE Step 2 CS CIS behavior list</th>
<th>Core competency</th>
<th>Medical Spanish core competencies and performance objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cercanía en la comunicación</td>
<td>Friendly communication</td>
<td>Fostering relationship</td>
<td>1, 2, 3, 4, 5</td>
<td>• Use appropriate register given the context of the interview (e.g., adult, child, family) and incorporating patient-centered and culturally responsive language and manner</td>
</tr>
<tr>
<td>Trato respetuoso</td>
<td>Respectful treatment</td>
<td>Fostering relationship</td>
<td>1, 2, 3, 4, 5</td>
<td>• Use appropriate register given the context of the interview (e.g., adult, child, family) and incorporating patient-centered and culturally responsive language and manner</td>
</tr>
<tr>
<td>Capacidad de escucha</td>
<td>Listening to my story</td>
<td>Fostering relationship, gathering information</td>
<td>1</td>
<td>• Obtain an accurate, age-appropriate complete medical history</td>
</tr>
<tr>
<td>Honestidad en la comunicación</td>
<td>Honest communication</td>
<td>Providing information</td>
<td>3, 4</td>
<td>• Orally communicate findings of the medical evaluation to the patient, adjusted for cultural, emotional, and literacy needs</td>
</tr>
<tr>
<td>Interés por mí como persona</td>
<td>Interest in me as a person</td>
<td>Fostering relationship</td>
<td>1, 2, 3, 4, 5</td>
<td>• Use appropriate register given the context of the interview (e.g., adult, child, family) and incorporating patient-centered and culturally responsive language and manner</td>
</tr>
<tr>
<td>Discusión sobre opciones/planes</td>
<td>Discussion of options/plans</td>
<td>Making decisions</td>
<td>4</td>
<td>• Orally communicate the treatment plan to the patient, adjusted for cultural, emotional, and literacy needs</td>
</tr>
<tr>
<td>Capacidad de animarme para hacer preguntas</td>
<td>Encouraging my questions</td>
<td>Fostering relationship, gathering information</td>
<td>3, 4, 5</td>
<td>• Assess patient comprehension of the information provided and address gaps in patient's knowledge</td>
</tr>
<tr>
<td>Claridad en la explicación</td>
<td>Providing clear explanation</td>
<td>Providing information</td>
<td>3</td>
<td>• Orally communicate findings of the medical evaluation to the patient, adjusted for cultural, emotional, and literacy needs</td>
</tr>
<tr>
<td>Exploración física</td>
<td>Physical examination</td>
<td>Fostering relationship, gathering information</td>
<td>1</td>
<td>• Engage and instruct the patient during a complete physical examination</td>
</tr>
<tr>
<td>Uso de vocabulario apropiado</td>
<td>Appropriate vocabulary</td>
<td>Providing information</td>
<td>2</td>
<td>• Successfully conduct the medical interview by asking appropriate questions that demonstrate evaluation of pertinent positives/negatives that sufficiently address differential diagnostic considerations for a given chief complaint</td>
</tr>
</tbody>
</table>

(Table continues)
pain, student asked what medications I take), were more likely to receive higher scores on the CAI scale. Faculty score on students’ case notes did not show a significant association with CAI results. Regarding other course performance metrics, student self-reported confidence on medical Spanish skills performance was positively correlated with CAI results for the shortness of breath case and overall (P < .01) but not significant for the abdominal pain case CAI scores alone. Student knowledge scores on a postassessment written examination did not correlate with CAI results. Higher student performance on the CAI was significantly associated with higher self-reported precourse Spanish levels (all P values < .05 for student CAI score on each of the clinical cases independently, and for both cases overall).

Consequences
We applied previously published Angoff minimum passing standards for the RUCIS scale17 to the CAI, including M2 level cutoff (defined as students with basic science training) and M4 level cutoff (defined as students with clinical training and patient exposure through clerkship rotations). If the M2 level cutoff score was applied to the students’ CAI performance, 12% would fail (compared with a 0% fail rate at the M2 level for the RUCIS scale reported in the prior study17). If the M4 level was applied, 29% of students would fail the CAI (compared with 10% in the RUCIS study). The course faculty provided learners with feedback, including individual CAI scores and an opportunity to watch their own recorded encounters. Students were asked to reflect upon and provide feedback about the SP encounters and CAI results. Overall, students stated that the SP encounters “allowed us to practice in real time” and gave them “a chance to try out medical Spanish skills and get valuable feedback.”21 Many student comments indicated a desire for an even greater amount of patient and/or SP encounters. Some suggested allowing more time in the encounter between obtaining the history and counseling the patient, so that they could gather their thoughts. In regard to how the SPs had perceived and scored their communication skills, students commented that they valued seeing how their use of Spanish contributed to developing relationships and better rapport with patients. Several expressed pride in their increased vocabulary and ability to properly phrase questions, but acknowledged that they “may still make some grammatical errors” such as “masculine/feminine agreements.” Students also reported greater awareness of when they require a medical interpreter, especially when the “conversation is too high stakes with concomitantly complex vocabulary, such as breaking bad news, or when there is a barrier to mutual understanding amongst the patient and student.”

Discussion
This validity study examined CIS assessment for Spanish SP encounters using the CAI scale, demonstrating supporting validity evidence in content, response process, internal structure, relations to other variables, and consequences. Adapting the RUCIS for use in medical Spanish cases involved not only translation but also an iterative process for appropriate linguistic and cultural adaptation. Feedback from reviewers allowed us to integrate perspectives from several Spanish-speaking nationalities, which enriched the linguistic variants and cultural concepts included in the final CAI.

CAI psychometric features
The CAI’s psychometric properties reveal several unique features compared with the 14-item English RUCIS scale on
which it was based. For example, the CAI’s learner variance accounted for 12% of total variance, which is almost 4 times greater than the learner variance reported for the RUCIS. Additionally, case specificity (how much the learners’ scores varied by case administered) was 2 times higher for the CAI than that reported for the RUCIS. CIS learner performance has been previously shown to be case-specific, but had never previously been evaluated in non-English cases. Our finding that the CAI assessment has higher case specificity than previously reported in English suggests that the nuanced contextual elements of individual patient scenarios may matter even more when learners are using limited or second-language skills in the encounter. This is consistent with literature regarding adult language learning that emphasizes the importance of using authentic materials to practice realistic communication skills for most effective language acquisition. It follows that the more authentic Spanish clinical cases are practiced and progressively assessed in medical Spanish programs (such as through formative feedback provided following role-play encounters, SP cases, or supervised patient interviews), the greater the utility for students’ overall communication skill advancement.

When we examined CAI scores’ relation to other medical Spanish performance objectives, we found that students who did not perform well on SP checklist items performed better on the CAI. This may be because students who spent more time during the encounter on CIS such as

### Table 3

**CAI Scale Item Difficulty and Reliability**

<table>
<thead>
<tr>
<th>CAI item by case</th>
<th>Item difficulty</th>
<th>Item discrimination (item-rest correlation)</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shortness of breath</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cercanía en la comunicación/Friendly communication</td>
<td>3.00 0.55 3–3</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Trato respetuoso/Respectful treatment</td>
<td>3.18 0.46 3–3</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td><strong>Abdominal pain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cercanía en la comunicación/Friendly communication</td>
<td>3.09 0.62 2–3</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Trato respetuoso/Respectful treatment</td>
<td>3.15 0.58 2–3</td>
<td>0.66</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CAI, Comunicación y Habilidades Interpersonales (Spanish for “Communication and Interpersonal Skills”); IQR, interquartile range; SD, standard deviation.

*Items are on a 1–4 scale, and overall scores are on percent scale.

*Item difficulty (mean and SD) is standardized to the traditional 0–1 measurement scale for the overall assessment tool.
building rapport and providing patient-centered explanations may not have spent as much time on factual information gathering, and is consistent with prior research.27 Relatedly, performance on the written knowledge assessment (which primarily evaluates medical Spanish vocabulary and grammar in health contexts) was not correlated with CAI scores, whereas higher student Spanish level was positively correlated with CAI performance. This finding suggests that general conversational skills in the target language are more relevant to the skills evaluated by the CAI than knowledge of technical vocabulary or grammar, consistent with the CAI’s intended emphasis on patient-centered communication rather than on medical vocabulary or linguistic correctness. These latter elements are also important to high-quality care of Spanish-speaking patients but may be better evaluated as separate medical Spanish domains by faculty raters rather than patients or SPs.

A multifaceted, multirater approach to medical Spanish assessment may help identify focused areas for improvement for individual learners, which understandably may differ by student language level.

Curricular implications

Implementing the CAI more broadly throughout medical Spanish programs may have important curricular implications. Using the CAI through formative and summative assessment and learning contexts would promote including patient-centered communication skills as part of curricular content and activities to adequately prepare students for the exam. This is in line with prior literature that recommends for medical language courses to move beyond solely teaching vocabulary or technical jargon and rather focus on effective strategies to communicate with patients.10,28 Course educational materials may also need to shift focus to authentic communication rather than vocabulary lists.29 Cases or other materials should incorporate social health issues such as immigration concerns, cultural practices, regional word choices, or health literacy variations. Cases that represent direct translations from English but do not represent the Spanish-speaking patient in an authentic fashion may not suffice to provide learners with realistic Spanish practice experiences.

Qualitative feedback from students suggests that receiving CAI scores following medical Spanish encounters may aid them in reflecting upon how their interpersonal skills are perceived by patients, identifying areas for improvement, and recognizing when they need a medical interpreter. Furthermore, to maximize the utility of the CAI as a formative tool for medical Spanish communication skills, the postcounter-guided reflection activity may help learners consider how they are perceived by a Spanish-speaking patient and create a plan to adjust their communication strategies or language study plan, as needed. Using the CAI as a formative tool not only can prepare learners for its use in summative assessment of skills at the end of the course or for language certification but also has the potential to help them better understand their communication strengths and limitations in Spanish. This approach supports language learning as a longitudinal process30 that benefits from multiple points of progressive assessment and individualized feedback rather than a single summative examination.

Implications for medical Spanish certification

The CAI’s variance, including learner and case specificity, suggests that up to 12 cases should be administered to achieve sufficient reliability. This finding is consistent with the number of cases used in USMLE Step 2 CS to test English skills1 and has important implications for potential medical Spanish certification exam development, in which the CAI could be applied as a summative tool for Spanish CIS assessment. Developing such a high-stakes examination should take into account the number of cases that need to be evaluated to fairly assess and certify the skills of students and physicians in non-English languages.

Further, it is critical to consider the appropriate cutoff ratings that will determine whether a candidate has passed or failed. Our data indicates that applying RUCIS Angoff cutoff ratings to the CAI would result in a higher fail rate for the CAI compared with the RUCIS. The majority of our student participants were at the M4 level, indicating that the M4 level cutoff scores would more appropriately reflect the expected level of CIS performance for most of our learners’ medical training. However, it is important to note that their prior exposure to medical training in Spanish is much more limited than their English training, bringing to question whether a different cutoff ought to be considered for Spanish skills. From a health equity perspective, we argue that lowering the cutoff standard in Spanish or other non-English languages would suggest that a lower communication quality standard is acceptable for LEP populations than for English-speaking patients. Medical Spanish assessment should ensure that certified bilingual clinicians are held to

---

Table 4

Factors Associated With Learner CAI Performance (n = 34)

<table>
<thead>
<tr>
<th>Case</th>
<th>Spanish CIS scale (% score)</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>0.71</td>
<td>0.16</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>0.72</td>
<td>0.18</td>
</tr>
<tr>
<td>Overall</td>
<td>0.72</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Abbreviations: CAI, Comunicación y Habilidades Interpersonales (Spanish for “Communication and Interpersonal Skills”); CIS, Communication and Interpersonal Skills; SD, standard deviation; SP, standardized patient.

<sup>a</sup>P value < .05.

<sup>b</sup>P value < .01.

<sup>c</sup>P value < .001.
equivalent communication standards for Spanish-speaking patients as they are for English speakers. Additional non punitive medical Spanish training may need to be supported for candidates who fail the exam to encourage progressive skills mastery and incentivize rather than hinder the development of more physicians with competent language skills.

Our study has several limitations. Two of the RUCIS items (physical examination and receptiveness to feedback) were adapted and refined for the CAI but were not tested in our study’s case scenarios. Additionally, we evaluated CAI scores using 2 cases only and found high case specificity. Future study should evaluate the CAI in a wider sampling of cases that include the 2 untested items. A standard setting meeting among medical Spanish educators to identify defensible cutoff points could be helpful in further evaluating pass/fail levels to determine minimal competence for medical Spanish CIS. Our study included students with variable Spanish levels, which is important to evaluate how well the tool works for students at different Spanish abilities, since use of limited Spanish skills in patient care is a common occurrence and speakers at all levels can benefit from feedback about how those skills are perceived by patients. Further study with a larger sampling of learners at multiple levels of Spanish may help to further inform recommendations for clinicians at specific levels. Additional research can also investigate whether the findings from this study are generalizable to other medical schools beyond the 2 in this study, other educational contexts (e.g., graduate or continuing medical education), CIS adaptations to other non-English languages, and CIS ratings for students when working with a professional interpreter during a language-discordant patient encounter.

Conclusions
CIS assessment is an important element in evaluating the skills of students or physicians who wish to provide independent patient care in non-English languages. The CAI represents the first tool specifically developed to evaluate future physicians’ patient-centered interpersonal communication skills in Spanish.

Acknowledgments: The authors wish to thank Maria Luisa Ortega Hernández, PhD, for her contributions to the CAI scale translation and Tania Gabriela Martínez and Noé Vázquez for their review of the CAI scale.

Funding/Support: None reported.

Other disclosures: P. Ortega receives author royalties from Saunders, Elsevier. The authors report no other conflicts of interest.

Ethical approval: This study was determined to be exempt by the institutional review boards of the University of Illinois at Chicago (Protocol #2019-0945) on September 13, 2019, and Northwestern University (Protocol #STC00212479) on November 17, 2020.

Disclaimers: The work presented here is that of the authors and does not necessarily represent the views or policy of the United Nations.

P. Ortega is clinical assistant professor, Departments of Medical Education and Emergency Medicine, University of Illinois College of Medicine, Chicago, Illinois; ORCID: http://orcid.org/0000-0002-5136-1805.

N.R. Moxon is a medical student, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

A.K. Chokshi is a medical student, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

C. Pérez-Cordón is a language specialist, Language and Communication Training Unit, United Nations Headquarters, New York, New York.

Y.S. Park is associate professor, Harvard Medical School, and director of health professions education research, Massachusetts General Hospital, Boston, Massachusetts; ORCID: https://orcid.org/0000-0001-8583-4335.

References


