Assessing Intercultural Competence

By: Loraine Phillips and Ryan McLawhon

Since implementing its International and Cultural Diversity (ICD) requirement in 2002, Texas A&M University has prioritized students' development of intercultural and global knowledge and competence. Students are required to complete six hours of coursework with an ICD designation, indicating significant content and activities designed to educate a "more pluralistic, diverse and globally aware populace" (Texas A&M University 2012, 19).

Developed in 2009 as part of the Academic Master Plan, the university's Teaching and Learning Roadmap established learning outcomes for undergraduate general education. These learning outcomes align with both the Association of American Colleges and Universities' (AAC&U's) Essential Learning Outcomes and the state of Texas's newly adopted Core Curriculum Objectives, to be implemented in fall 2014. The university specifies that students will "demonstrate social, cultural, and global competence, including the ability to

- live and work effectively in a diverse and global society;
- articulate the value of a diverse and global perspective;
- recognize diverse economic, political, cultural and religious opinions and practices." (Texas A&M University 2012, 22)

To assess student achievement in these areas, Texas A&M's Office of Institutional Assessment (OIA) asked faculty members teaching the ten most popular ICD courses to submit complete sets of student work that had been assigned with the intention of evaluating students' social, cultural, and global competence. As a result of this inquiry, the assessment office discovered that faculty in these high-enrollment, lecture-driven courses relied primarily on objective exams to evaluate student performance. Since these exams were not an ideal tool for assessing the nuances of social, cultural, and global competence, the office sought another avenue for harvesting student work.

Pilot Project

Starting over, the OIA drew on the resource of existing faculty assessment liaisons from each college. **These liaisons contacted their department chairs to identify faculty members who were assigning work related to intercultural, global, and/or diversity issues.** Communication with those faculty members yielded seventy-four student papers from two of Texas A&M's ten colleges (the College of Liberal Arts and the College of Education and Human Development). After redacting any personally identifiable information, the assessment office coded the papers for analysis.

The assessment liaisons decided **to use AAC&U's Valid Assessment of Learning in Undergraduate Education (VALUE) rubric on Intercultural Knowledge and Competence (Rhodes 2010) to evaluate the papers for evidence of social, cultural, and global**

competence. This rubric describes four levels of student performance related to a variety of different criteria, including cultural self-awareness, knowledge of cultural worldview frameworks, empathy, communication, curiosity, and openness. To test the rubric, staff and faculty volunteers met during summer 2011 with the assistant vice president for Global Programs Support, who had extensively studied related assessment literature. During a one-day event, the group calibrated their assessments by discussing their expectations for applying the rubric criteria, identifying an "anchor" paper for each assignment, and scoring those papers together to develop consensus about an approach. Two participants then scored each remaining paper, with a third participant scoring any papers where the first two scores diverged widely. OIA staff calculated interrater agreement scores to identify places where discussion and recalibration might be beneficial. Over the course of the eight-hour workday, each participant scored approximately nine papers. Following the scoring session, scorers participated in a focus group discussion to evaluate the rubric's usability, its applicability to student work, its calibration and scoring, and the associated workload.

The rubric was well received by participants, who had several recommendations for future administrations. First, the group recommended removing the criterion for "Skills: Verbal and nonverbal communication," because these skills are difficult to evaluate via written assignments. Next, participants rejected the possibility of establishing a minimum page requirement for papers to be evaluated, since even short papers could address all rubric criteria. Participants reinforced the importance of calibrating each individual assignment type before scoring student papers and discussed the need for diversity (of ethnicity, gender, department/college affiliation, etc.) among faculty scorers. The group agreed that the overall workload was acceptable and the time allocated for calibration and scoring was appropriate.

Intercultural Competence Project

Based on pilot project findings, Texas A&M launched the Intercultural Competence Project (ICP) in summer 2012. Assessment liaisons extended e-mail invitations to faculty asking for student work dealing with international, global, or diversity issues, and faculty submitted nearly two hundred and fifty papers in response. As in the pilot project, the assessment team redacted personally identifiable information and coded the papers for analysis. For this administration, the team also obtained student demographic information.

The associate dean of diversity in the College of Liberal Arts joined the assistant vice president for Global Programs Support in leading the calibration process. Using a modified version of the rubric that excluded the verbal and nonverbal communication skills criterion, faculty from four different colleges scored the submitted papers.

Following the scoring exercise, OIA staff compared mean scores based on gender, ethnicity, and classification (for example, junior). Although the scores showed no statistically significant differences for gender and ethnicity, seniors scored better on average for each criterion. The OIA created department-level reports for each participating unit, comparing each department's scores to the overall average. These reports should help spark conversation about opportunities to enhance pedagogy and the curriculum.

Areas for Improvement

In both the pilot project and the ICP, faculty participants emphasized the importance of seeing the original writing prompt for each group of papers. In future administrations, the OIA will obtain these prompts and summarize them for faculty scorers. Several scorers were uncomfortable assigning scores for particular criteria that faculty members may not have emphasized in the course.

In both administrations, scorers reached the lowest level of agreement about how to score the "Attitudes: Curiosity" section of the rubric, which they found difficult to interpret. The OIA plans to study related assessment theories to provide deeper explanations for this criterion. The lower agreement scores for this section could also have been affected by some scorers' tendency to evaluate students' curiosity in comparison to the perceived curiosity levels of typical Texas A&M students, rather than in direct relationship to the rubric's criteria. Faculty members also disagreed about whether specific actions (for example, studying abroad) conclusively indicated high levels of curiosity.

Future administrations of the ICP will target the eight colleges not included to date. Although the racial composition of the first ICP sample was representative of Texas A&M's student population (approximately 70 percent white), the OIA has also begun identifying strategies to obtain more work by students from underrepresented groups.

With support from the OIA, assessment liaisons will present project results, including processes and methodologies, to college leaders. By including as many stakeholders as possible and adjusting processes according to their feedback, the assessment office aims to yield more relevant and actionable results. Ultimately, the office aims to establish the ICP as one of many measures of student learning designed to inform curricular and pedagogical improvements at Texas A&M.

Planning a Scoring Day

To institutions considering planning a scoring day using rubrics, we recommend the following steps:

1. Find, gather, and prepare student papers.
2. Recruit subject matter experts as calibrators.
3. Plan paper scoring order with calibrators.
4. Estimate time needed to score assignment types and block day accordingly.
5. Invite scorers.
6. Designate staff to calculate real-time interrater agreement and manage paper distribution.
7. Order food!

—Loraine Phillips and Ryan McLawhon

References

Association of American Colleges and Universities. 2011. *The LEAP Vision for Learning: Outcomes, Practices, Impact, and Employers' Views*. Washington, DC: Association of American Colleges and Universities.

Rhodes, Terrel L., editor. 2010. *Assessing Outcomes and Improving Achievement: Tips and Tools for Using Rubrics.* Washington, DC: Association of American Colleges and Universities.

Texas A&M University. 2012. *Texas A&M University 2012–2013 Undergraduate Catalog.* http://catalog.tamu.edu/pdfs/12-13_UG_Catalog.pdf.