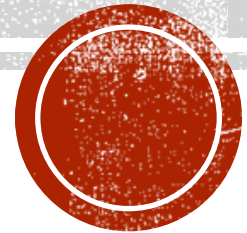


GENERAL EDUCATION ASSESSMENT: RUBRIC CALIBRATION AND SCORING

Presented for the Learning Assessment Committee

February 20, 2015

Eileen DeLuca, Joe van Gaalen, and Amy Trogan



TIME FRAME FOR RATERS

- **Session 1 (2 hours):**
 - **Rubric Calibration**
 - **Score approximately 30 artifacts**
- **Session 2 (2 hours):**
 - **Score approximately 45 artifacts**



RUBRIC CALIBRATION

“Calibration is a process in which those using a rubric or rubrics gather to achieve a common understanding of how the rubric was designed and how it ought to be applied” (Rhodes and Findley, 2013, p. 23).



CALIBRATION PROCEDURE (ADAPTED FROM RHODES AND FINDLEY, 2013).

- 1. Close reading
- 2. Discuss ambiguities
- 3. Agree on interpretation of rubric language
- 4. Read and Score sample work
- 5. Review scores to determine a consensus



CLOSE READING, DISCUSSION, AGREE ON INTERPRETATION

- Read through rubrics (front page and back page)
- Identify ambiguities in phrasing or wording (use highlighter)
- Discuss questions raised and come to agreement on how rubric's language should be interpreted for the purpose of scoring.
- *Note: The calibration session is not the time to make changes to the rubric.



READ AND SCORE SAMPLE WORK

- **Follow this procedure for 1-2 sample artifacts.**
- Score one criterion (row) at a time.
- After each row has been scored, review to determine the degree to which consensus has been reached.
- Ask faculty to gave an artifact a particular score (e.g. “2”) should be asked to identify places in the sample work to support that decision. Then a participant who chose a distinct score (e.g. “4”) should be asked to provide a similar rationale.
- During calibration participants have the opportunity to change scores.
- The goal of calibration is not to have a unanimous agreement for one single score. Rather to identify two scores around which the majority clusters.
- Score at least two artifacts in the manner described above.



SCORING SESSION

- Following the calibration session, teams will be provided copies of their artifacts **(electronic?)**
- Raters will score individually (no discussion with team member).
- While scoring, raters will **(enter scores in spreadsheet? Write on actual rubric copies)**
- During scoring, raters will make notes on the “Qualitative Feedback sheet” to discuss strengths and challenges of rubric.



POST-SCORING

- Depending on time, team leaders can lead a discussion about the process.
 - This is not a time to compare scores, but rather to talk about whether or not the rubric is a good fit to measure achievement.
 - Also, it is a time to discuss whether or not the rubric would require revision before implementation.
 - Team leaders can take notes on discussion themes.
- Team leaders will collect scores and qualitative feedback sheets and turn it to Crystal Revak.



ANALYSIS

- Eileen will analyze qualitative feedback for major themes
 - Provide a report (by rubric) of strengths.
 - Provide a report of any suggested revisions.
- Joe will analyze rubric scores
 - Provide overall mean scores for each criterion on each rubric.
 - Provide disaggregated analyses (e.g. overall mean scores of courses traditionally taken in 1st year vs. 2nd year, overall mean scores by discipline, etc.)
 - Provide inter-rater reliability (overall and by rubric dimension) for each rubric



CLOSING THE LOOP

- Analyses will be disseminated to Learning Assessment Committee for review and conclusions.
- Assessment Coordinators will report initial results during department meetings.
- Samples of finding shared in **Assessment 101 Part II?**
- Tentative: Assessment report for faculty in **May?**



MODEL CALIBRATION SESSION

- Facilitators: Amy, Eileen, Joe



REFERENCES

Rhodes, T.L. & Finley, A. (2013) Using the value rubrics for improvement of learning and authentic assessment. Association of American Colleges and Universities.

