Robust reliability of rubrics? Examination of faculty-faculty and faculty-student rubric ratings

Nancy E. Frye, Michele M. Dornisch, General Education Outcomes Assessment Task Force*

Nancy.Frye@liu.edu

Long Island University-Post

Michele.Dornisch@liu.edu

*Glynis Pereyra, John Lutz, Wendy Ryden, Kathy Isoldi, Carol Boyer, Andrea McLoughlin, SeungYeon Lee, Jeong-eun Rhee, Eduardo Rivera, Yen Yen Woo

Thanks to all those who completed the rubrics and had students complete assignments in their classes

Use of Rubrics

- Increasingly, rubrics are being used to assess specific areas of student strengths and weaknesses
- For instance, LEAP rubrics and modifications of these are used to assess student skills (e.g., critical thinking)
- These rubrics are used both to examine where students have strengths and weaknesses, and to give feedback to students

Meaningful Feedback

- For the feedback to be meaningful to students, students must interpret the rubrics in the same way as faculty
- This can be tested by comparing student self-ratings to faculty ratings
  - Such comparison suggests that, while students see their overall level as higher than faculty do, there is agreement on specific areas of strengths and weaknesses (Frye & Dornisch, 2016)

Self-ratings?

- When student self-ratings are used, they may overestimate their abilities, due to self-serving bias, or due to a concern that their self-ratings may impact their grade
- To optimally assess students' perceptions of rubrics and how they agree with faculty perceptions, one should examine the perceptions of students whose views are relatively objective
  - i.e., not the students who completed the work

How to assess agreement?

- Two ways to address:

1. Agreement in ratings on each dimension of the rubric?
   (Agreement in absolute level of achievement)
2. Consistency in perception of relative areas of strength and weakness?
   (Agreement in relative level of achievement)

Recruitment and Classes

- Data were collected during the 2015-2016 Academic Year
- Participation was open to any class with undergraduates who were primarily seniors
  - Before the school year, faculty were presented with a template of a critical thinking assignment and asked to incorporate it in their class
    - [Three problems in the content area of the class listed] Consider one of the three problems listed above: Complete a critical analysis of this problem. Include, among other aspects of your analysis, an explanation of the problem and its complexities as well as solutions/actions that you determine could or should be taken based on your analysis. Also consider the implications and consequences of the solution(s)/ action(s). Be sure to properly cite sources you use.

Classes and study design

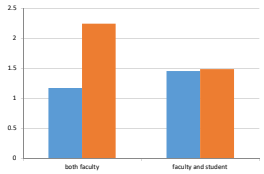
- 105 papers were collected from seniors across 8 classes (history, social work, sociology, economics, nutrition, criminal justice, and two English classes)
- After the semester, faculty and students rated the papers using a modified version of the LEAP rubric for critical thinking and for integrative reasoning
- 92 papers were assessed by 2 coders
  - 76 by one faculty member and one student
  - 16 by two faculty members

Analyses

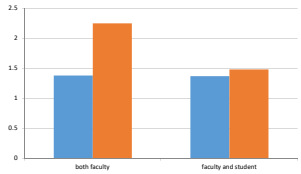
- Comparison of level
  - One 2 (coder) X 2 (type of pair) ANOVA was used for each row of the rubric
- Comparison of pattern of ratings
  - Multilevel modeling was used to examine patterns of agreement across rows of the rubric
  - Type of coding pair was entered in level 2, to see if patterns of agreement varied by type of coding pair

Findings: Rating levels

- Significant coder X type of pair interaction for 3 of 7 rows of rubric
  - Questioning experts:

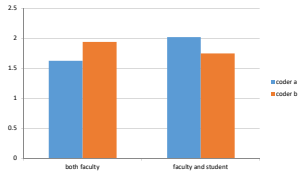


Connections:



Findings: Rating levels

- The two previous interactions suggest greater agreement between faculty-student pairs than between faculty-faculty pairs
- An additional interaction, with respect to position, shows a different pattern:



Findings: Relative patterns of strengths and weaknesses

- Positive and significant association between ratings of two coders
- No significant difference in amount of agreement between student-faculty and faculty-faculty pairs

Supplemental analyses

- Examined whether type of student paper mattered
- GPA of student writing the paper was entered in level 2 of the multilevel model

→ Greater agreement among coders (regardless of type of coding pair) when the writer of the paper had a higher GPA

Rubric

Dimension	Highest level descriptor
Explanation	Issue/problem to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding.
Evidence	Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis.
Question	Viewpoints of experts are questioned thoroughly.
Context	Thoroughly (systematically and methodically) analyzes own and others' assumptions and carefully evaluates the relevance of contexts when presenting a position.
Position	Specific position (perspective, thesis/hypothesis) is imaginative, taking into account the complexities of an issue. Limits of position (perspective, thesis/hypothesis) are acknowledged. Others' points of view are synthesized within position (perspective, thesis/hypothesis).
Implications	Conclusions and related outcomes (consequences and implications) are logical and reflect student's informed evaluation and ability to place evidence and perspectives discussed in priority order.
Connections	Independently creates wholes out of multiple parts (synthesizes) or draws conclusions by combining examples, facts, or theories from more than one field of study or perspective.

Implications

- Do students and faculty use rubrics similarly?
  - Greater similarity when the student rating a paper is not the student who wrote the paper
  - There may be concerns about faculty using rubrics similarly
    - Critical thinking may mean different things in different disciplines

    → How best to get an overall understanding of students' critical thinking across disciplines in an institution?