

Hello Everyone!

This is Shivaji. Just thought I'd share some of my thoughts with you regarding evaluation and assessment. You have just finished both - or are finishing - for the Spring Cycle Of 2017.

As you all know, in the context of our College, evaluation is the faculty member's individual evaluation of his/her student's work, and assessment is the Committee's ratification of the faculty member's evaluation of the same student. Today, in this letter, I'd like to concentrate on what is known in contemporary academia as "rating," its validity and reliability. After all, the higher education in the entire country is agog with many vexing questions regarding assessment, validity and reliability being two of the principal ones¹. People often tend to use the term rating interchangeably with grading. Grading and Rating are similar but not the same. Grading may be defined as an evaluative matrix imputed to the student's overall quality of work whether in an exercise or a course. Rating is the judgment of how well a learning item (knowledge or skill) has been mastered (Pendlebury, 2013). In Boricua College's scale of rating SM (Superior Mastery) is the best, U (Unsatisfactory) is the worst. Both grading and rating are done based on a set of *rubrics*, i.e., a set of qualitative criteria indicating mental, affective or psychomotor processes. Here, this rating system is used not only as an overall indication of competency, but the competency or the knowledge item is further subdivided into sub criteria and rated. As we explained in a recent self-study:

[The faculty] start with evaluation of the student's general disposition toward his/her academic responsibility indicated by attendance, punctuality, participation and discussion, presentation of work, and general ability with the cognitive and affective skills. From there, the rubrics evaluate each of the five components of the intellectual competency cluster, as well as the Affective skills

¹ See for example Kevin Gannon's piece, "Stuck in the Assessment Swamp?" that came out this week in *Vitae*, a supplement of *The Chronicle of Higher Education*. May 1, 2017

cluster, evaluating the overall and the component sub-skills. The assessment form also contains an evaluation of the student's knowledge he/she has learned in each course. (*Inquiry Brief to TEAC – 2013*)

The problem of validity and reliability comes from centuries long traditional grading practices, where faculty's grading has been beyond questioning. Those in charge of determining accreditations do not anymore consider grades as "primary." There is a high degree of subjectivity in such grading, they say. Therefore, while assessing competency-based student learning, "rating" is regarded as more reliable than traditional grading (Banta, 2017). What is the difference?

While traditional grading simply indicates a student's overall achievement, there is little or no indication whether the student has met the course objectives, or for that matter, the Department's goals and objectives. To the traditional faculty member, the grade is indication enough - for everything. Boricua College's assessment method measures smaller and smaller criteria of the learning item, rating each sub skill *according to the objectives of the course, and according to the written and/or **oral** evidence in the student's work*. This makes it possible for those who are analyzing student learning to determine, not only the overall strengths and concerns about the student, but also *in what sub-items exactly* the student needs to improve, and similarly, where the student's specific strengths lie. In other words, rating is more valid than grading. They are valid:

Because the rubrics developed...have been based for all three domains of the Bloom Taxonomy, which itself has found general acceptance in the educational community for over 50 years, the rubrics can be considered to have *construct validity*. (*Inquiry Brief, 2013*)

Similarly, reliability. We explain in the same document that reliability entails:

assurance that an assessment will yield comparable results under similar circumstances over and over again, regardless of who administers it. In terms of this self-study,

reliability is a measure of assurance that the rubrics...can be assumed to provide comparable results year after year.

From the evidence of our external accreditors, we are succeeding. In the words of John Ettl, the most recent Team Leader of the Middle States visiting the College:

To date, the strongest assessment efforts have been found in the development of the learning assessment framework, evidenced by the implementation of an ongoing systemic assessment cycle for all programs with collaborations/input of faculty and administrators; establishment and ongoing assessment of program-specific expectations (direct measures) for student learning in alignment to institutional learning goals; development and continuous revisions with faculty input of syllabus and standardized rubrics to evaluate intellectual, affective, and psychomotor domains; program assessment reports developed by faculty, submitted to the president and shared in faculty department meetings, colloquia and annual three-week development workshops to review data (baseline and comparative findings) and use it for program and course improvements. (Middle States Report, 2016)

But the dialogue in the country goes on. Gannon, in the above-mentioned article published by the *Chronicle* says what needs to be measured is the process of learning, not merely products: “‘outcomes’ *is* the process.” (italics inserted) True; but he never explains how process may be assessed. Boricua’s rubrics rating intellectual, affective and psychomotor skills are one of the best attempts to do just that.

In a very interesting prolonged discussion on Assessment of Student Learning, sponsored by LinkedIn², I had proposed to Professor David Rehm (who had earlier served as a Team Leader of the Middle States to our College) that the breaking down of skill clusters into sub-skills, and linking them to the course goals and objectives, and thence to program goals goes a long way to giving an actual, real-time assessment of whether the student or, by aggregate, the Department, has successfully met the program goals. He agreed.

² LinkedIn is an online professional website which sponsors discussions among professionals all over the world on various topics of almost every single recognized profession. Membership is free, up to a point. As LinkedIn member, I participate regularly in several professional discussion groups, one of the being Assessment of Student Learning.

So then, what am I saying? I think, as possibly the *only* competency based liberal arts institution in the U.S., we have a lot to contribute to American higher education if we evaluate and assess our students' work *meticulously*, with an eye for maximum details. The more minutely we rate a learning item, the better the College's chances to improve not only the student, but also the Department, and, ultimately determine whether the Mission of the College is being fulfilled.

Of course, we have a lot of work to do! For one thing, we do not really have a proper set of rubrics in Theoretical and Cultural Studies. Consequently, these courses are still graded in the traditional sense. We should create rubrics for these courses. Also, we have the same rubrics for all levels of cognitive, affective and psychomotor skills development from the basic Generic Studies right up to the Master's. Should we adapt those rubrics according to levels?

These are important questions, and need discussion.

Thank You.

BIBLIOGRAPHY

Banta, C. P. (2017). *Assessment Essentials*. San Francisco: Jossey-Bass.

Boricua. (2013). *Inquiry Brief of the Education Department Submitted to TEAC*. New York: Boricua College.

Boricua. (2016). *Middle States Self Study 2016*. New York City: Boricua College.

Gannon, K. (2017, May 1). Stuck in the Assessment Ramp? *Vitae*, pp. 1-7.

Pendlebury, G. B. (2013). *Assessing Student Learning in Higher Education*. London: Routledge.