

Using Rubrics with Technology for Formative Assessment of Faculty Teaching Effectiveness

Adele Weiner, Ph.D., L.C.S.W.
Metropolitan College of New York
United States
aweiner@mcny.edu

Abstract: Studies have shown that the use of rubrics increases student learning and achievement. This presentation demonstrates how rubrics, generated for individual student evaluation in the everyday assessment of learning, can also be used for understanding class performance and inform teaching strategies. Technology is used in the creation and grading of rubrics, which allows them to be easily coded and collected for immediate computer generated data visualization and statistical analysis. This type of assessment can be used across multiple classes by faculty to provide feedback for formative teaching evaluation, continuous improvement, course development, program self-study and accreditation processes.

A rubric is an explicit set of criteria used for assessing a particular type of work or performance. A rubric usually also includes levels of potential achievement for each criterion, and sometimes also includes work or performance samples that typify each of those levels. (Ehrmann, 2011)

Multiple studies have shown that the use of rubrics increase student learning and achievement. In recent years, rubrics have been successfully incorporated into college settings (Vandenberg, Stollak, McKeag, and Obermann, 2010; Lencho, Longrie, and Friedman, 2009; Miller, 2005). *Formative assessment is ongoing assessment that is intended to improve an individual student's performance* (Middle States Commission on Higher Education, 2008, p. 34). Rubrics are used for formative assessments as part of instructional design to provide crucial feedback for teachers and students (Garrison and Ehringhaus, 2007).

Faculty use other tools to assess their own teaching effectiveness. These include peer observations, course evaluations, self-evaluation reports, teaching journals, and portfolios. This paper proposes that faculty can use the rubrics prepared for student assignments to both assess their own teaching and as a basis for research on teaching practices. Analysis of assignment rubrics at the class level can provide faculty with formative assessment to improve their teaching.

A Rubric for Assessing Written Assignments

Writing is a skill that all students need to develop along with other communication skills. Beginning in the Fall of 2006, this author began to create a generic writing rubric to provide students with guidance in preparing their written assignments. It contained seven criteria that were to be evaluated: 1) organization of the paper, 2) the quality of the information, 3) all assignment questions answered, 4) relevant references, 5) application of theoretical/conceptual constructs, 6) grammar, and 7) APA citation style. Each of these criteria included four levels of achievement (Excellent, Good, Fair and Poor) that were clearly described on a rubric grid (See Figure 1). Plagiarism was added at a later date as it became obvious, through this evaluation process, that plagiarism and using proper citations were not synonymous.

These grids were distributed in class and discussed prior to the first papers in a course to help students understand faculty expectations and the standards that would be used to grade their papers. A copy of the rubric was also posted to the class Learning Management System (Blackboard and then Moodle) course web site. When grading a paper, the faculty member marked a copy of the rubric and returned it to the student with the hand-graded paper. The copies of the rubric sheets were maintained for later manual entry into a spreadsheet for data analysis.

	Poor	Fair	Good	Excellent
Organization	The information appears to be disorganized.	Information is organized, but paragraphs are not well constructed.	Information is organized with well constructed paragraphs.	Information is very organized with well constructed paragraphs and subheadings.
Quality of Information	The Information has little or nothing to do with the main topic.	Information clearly relates to the main topic. No details and/or examples are given.	Information clearly relates to the main topic. It provides 1-2 supporting details and/or examples.	Information clearly relates to the main topic. It includes several supporting details and/or examples.
Answers Assignment Questions	The paper does not answer the questions of the assignment.	More than one of the questions of the assignment are not answered.	Most of the questions of the assignment are clearly answered.	All questions of the assignment are clearly answered.
Relevant References	There are no relevant references to course materials and/or other resources.	There are minimal relevant references to course materials and/or other resources.	There are some relevant references to course materials and/or other resources.	There are adequate relevant references to course materials and/or other resources.
Application of Theoretical/Conceptual Constructs	There is no application of theoretical/conceptual concepts to the material discussed in the paper.	There is minimal application of theoretical/conceptual concepts to the material discussed in the paper.	There is some application of theoretical/conceptual concepts to the material discussed in the paper.	There is adequate application of theoretical/conceptual concepts to the material discussed in the paper.
Grammar	Many grammatical, spelling or punctuation errors.	A few grammatical, spelling or punctuation errors.	Almost no grammatical, spelling or punctuation errors.	No grammatical, spelling or punctuation errors.
APA Citations	Some sources are not accurately documented.	All sources (information and graphics) are accurately documented, but many are not in the desired format.	All sources (information and graphics) are accurately documented, but a few are not in the desired format.	All sources (information and graphics) are accurately documented in the desired format.
Plagiarism	Large portions of this paper are copied from other sources without proper citations. See your turnitin originality report.	Some portions of this paper are copied from other sources without proper citations. See your turnitin originality report.	Minimal portions of this paper are copied from other sources without proper citations. See your turnitin originality report.	No portions of this paper are copied from other sources without proper citations. See your turnitin originality report.

Figure 1: Written Assignment Rubric

Technology makes the process of marking rubrics and inputting data for analysis much easier. Beginning in the Spring 2008, the rubric was attached to electronically graded assignments using Turnitin.com (<http://www.turnitin.com>). Turnitin is an internet service that allows the submission and grading of papers online, use of rubrics, and checks for potential plagiarism. This program provides a permanent electronic record of the rubric and generates a report of individual scores and class averages (See Figure 2). The matrix of individual and class averages can easily be exported to excel or statistical software. This paper will discuss how this grouped, class data can be used to provide class feedback and also on the use of data analysis of course average scores for written assignments as a method for assessing teaching.

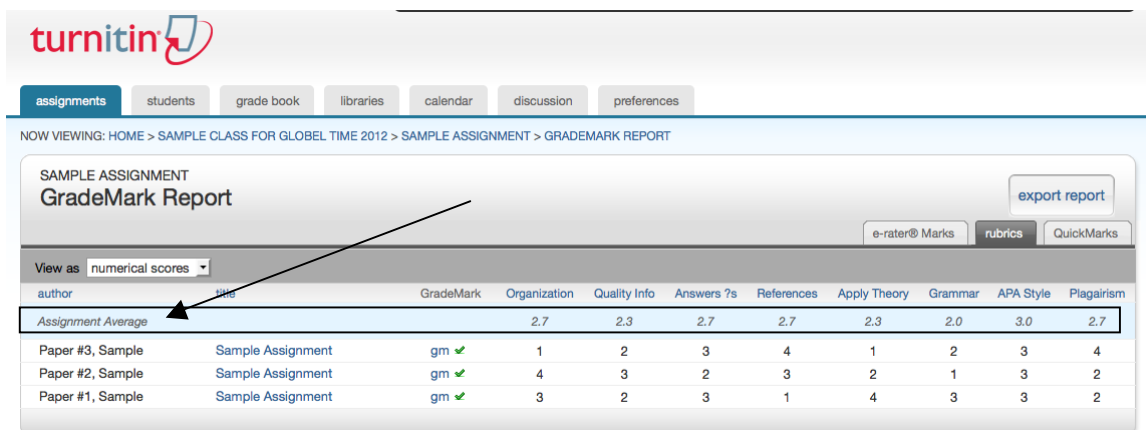


Figure 2: Turnitin Rubric Statistics

Assignment Evaluation

Faculty members are able to use the class averages generated by Turnitin to review class performance immediately following the grading of an assignment. It is possible for faculty to use Turnitin to export the class values on an assignment to an Excel spreadsheet or statistical software to generate a chart for rubric averages and/or frequencies (See Figure 3). This provides timely feedback to the class and allows faculty to assess whether the class as a whole or individual members of the class need additional help in meeting the criteria for written assignments. These charts are shared and discussed in class allowing both faculty and students to gain a broader understanding of the requirements and issues involved in achieving them. A secondary function of these charts is help students interpret data presented in this manner. Faculty can also use these charts as a method to assess, and if necessary adjust, teaching of required skills.

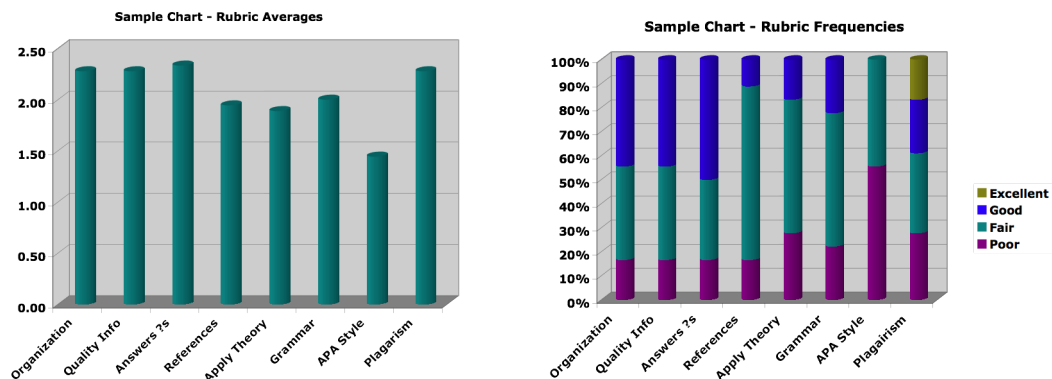


Figure 3: Sample Charts – Class Rubric Averages and Class Rubric Frequencies

A Model for Ongoing Formative Teaching Evaluation

From Spring 2008 through Spring 2011, undergraduate students, in 20 courses, were required to submit papers through Turnitin.com. All assignments (N=45) used the same rubric. Papers that were highly plagiarized were returned to students for revision and not marked using the rubric. Only the student's first attempt at each assignment is included here.

The purpose here is *not* to provide information about undergraduate performance in writing papers, but rather to demonstrate how authentic measures such as rubrics, generated for individual student evaluation in the everyday assessment of learning, can also be used for understanding class performance and inform teaching strategies.

Each level of the measurement criteria was awarded points: 4 for excellent (A), 3 for good (B), 2 for fair (C) and 1 for poor performance (D). These also correspond to standard points used in computing a student's grade point average (GPA) allowing easy conversion to letter grades. In some cases students received "0" points for no evidence of the category, which in most cases meant that the student failed to submit the paper or there was no adequate information for the faculty to assess the criteria. For example, many students failed to use any citations so faculty members were unable to determine if students had used APA style properly. Papers were graded online using Turnitin.com and the rubric was attached to graded papers. Students would go into their Turnitin accounts to retrieve graded papers with faculty comments and rubrics.

Mean class scores, for each assignment, were entered into a PASW/SPSS database. This analysis does not compare individual student scores, but the assignment mean scores for a class. The first analysis computed the average score for each rubric category (See Table 1).

	N	Minimum	Maximum	Mean	Std. Deviation
Organization	45	1.00	3.75	2.4098	.52925
Quality of Information	45	1.71	3.50	2.6100	.35452
Answers Questions	45	1.00	3.33	2.3874	.49316
Relevant References	45	1.00	3.29	2.0071	.68663
Applies Theory	45	1.17	3.00	2.0816	.54450
Grammar	45	1.43	3.50	2.5300	.44081
APA Style	45	1.00	2.75	1.6176	.51429
Plagiarism	9	2.27	3.60	2.8778	.38220
Valid N (listwise)	9				

Table 1: Descriptive Statistics

For all the categories, except APA style, it appears that students do “fair” work, or earn the equivalent of “C” to “C+” grades. The students’ poorest performance, with the lowest mean score, was in the use of APA style (mean = 1.6), which is a rather concrete task. The best performance was in not plagiarizing (mean = 2.9), which has been a faculty goal across the college.

The mean class scores provide ongoing formative assessment of class performance in meeting course standards and as a feedback mechanism for class discussions. These findings led to faculty development of multiple strategies for improving the use of the APA format, including discussion of technology resources such as the cite generation functions of the online library journal databases, online APA tutorials, and online citation services. Links to all of these resources are provided through the Moodle course site. Future analyses will help determine if these are successful.

A second analysis was performed to see if the performance on the criteria correlated with the other measures. (See Table 2). A number of correlations were found which could be used for evaluation of student learning and teaching strategies.

In response to this correlation analysis the faculty chose to focus on having the students learn to answer all of the assignment questions, another concrete skill. A strategy was developed that suggested they copy the assignment into a blank word processing document and then write their answers under each question. Since the syllabus and assignments were online, in Moodle, they simply had to copy and paste the assignment questions. As part of their final editing they were to turn the questions into topic statements for paragraphs. Discussions also included how to use relevant references and apply theory in answering the questions. In order to assist students with these tasks, the faculty revised the directions for written assignments to include specific statements as to the readings that should be used in answering the questions. Some of the questions also identified specific theoretical concepts that students should discuss in their completed assignments.

Conclusion

More important than the actual findings of this particular rubric analysis, is the evaluation process, which *uses computer generated aggregate data produced as a normal function of the teaching process*. It does not require faculty to create another tool for self-assessment. The mean class scores provide ongoing formative assessment of class performance in meeting course standards and as a feedback mechanism for class discussions. Technology reduces the amount of time and energy that has to be expended to produce statistical findings and can generate data visualizations that help in understanding patterns. If the rubric truly represents the kinds of expectations faculty have for student performance, faculty can identify deficiencies, design teaching strategies to address such deficits, and assess if these strategies are successful in improving performance. Documentation of this process can be useful in

demonstrating teaching excellence in portfolios and other periodic reviews. Since the data is grouped, the potential of an idiosyncratic effect of an individual class is reduced.

		Organization	Quality of Information	Answers Questions	Relevant References	Applies Theory	Grammar	APA Style	Plagiarism
Organization	Pearson Correlation		.485**	.317*	-0.002	0.221	.504**	0.24	0.313
	Sig. (2-tailed)		0.001	0.034	0.991	0.144	0	0.112	0.413
	N		45	45	45	45	45	45	9
Quality of Information	Pearson Correlation	.485**		.486**	.426**	.523**	0.211	.387**	0.354
	Sig. (2-tailed)	0.001		0.001	0.004	0	0.164	0.009	0.351
	N	45	4	45	45	45	45	45	9
Answers Questions	Pearson Correlation	.317*	.486**		.474**	.648**	0.062	0.232	0.407
	Sig. (2-tailed)	0.034	0.001		0.001	0	0.686	0.125	0.277
	N	45	45		45	45	45	45	9
Relevant References	Pearson Correlation	-0.002	.426**	.474**		.737**	0.179	.358*	0.066
	Sig. (2-tailed)	0.991	0.004	0.001		0	0.24	0.016	0.866
	N	45	45	45		45	45	45	9
Applies Theory	Pearson Correlation	0.221	.523**	.648**	.737**		1	0.229	0.374
	Sig. (2-tailed)	0.144	0	0	0		0.697	0.13	0.321
	N	45	45	45	45		45	45	9
Grammar	Pearson Correlation	.504**	0.211	0.062	0.179	0.06		1	-
	Sig. (2-tailed)	0	0.164	0.686	0.24	0.697		0.084	0.056
	N	45	45	45	45	45		45	9
APA Style	Pearson Correlation	0.24	.387**	0.232	.358*	0.229	0.26		1
	Sig. (2-tailed)	0.112	0.009	0.125	0.016	0.13	0.084		0.471
	N	45	45	45	45	45	45		9
Plagiarism	Pearson Correlation	0.313	0.354	0.407	0.066	0.374	-0.056	0.471	
	Sig. (2-tailed)	0.413	0.351	0.277	0.866	0.321	0.887	0.2	
	N	9	9	9	9	9	9	9	

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Table 2: Correlation Analysis

If some additional demographic information is added to the database (class name, course level, assignment number, semester, etc.) faculty can measure change in specific skills areas over time. They can also look at changes between assignments in the same class, original and revised assignments, lower level and upper level classes and different sections of the same course. This type of technological, quantitative course assessment can be used across multiple classes by various faculty to provide feedback for formative teaching evaluation, continuous improvement, course development, program self-study, and accreditation processes.

References

Ehrmann, S.C. (2011). Rubrics: Definition, tools, examples, references. *The Flashlight Evaluation Handbook*. The TLT Group (Teaching, Learning and Technology), Retrieved on January 1, 2012 from <http://www.tltgroup.org/resources/flashlight/rubrics.htm>

Garrison, C., & Ehringhaus, M. (2007). Formative and summative assessments in the classroom. Retrieved on January 1, 2012 from <http://www.nmsa.org/Publications/WebExclusive/Assessment/tabid/1120/Default.aspx>.

Lencho, M.W., Longrie, M.J., & Friedman, S. J. (2009). Where learning and assessment meet. *Assessment Update*, 21(2), 5-7.

Middle States Commission on Higher Education (2008). *Student Learning Assessment: Options and Resources*, Philadelphia, PA: Middle States Commission on Higher Education.

Miller, R. (2005). Integrative learning and assessment. *Peer Review*, 7(4), 11-14.

Vandenberg, A., Stollak, M., McKeag, L., & Obermann, D.. (2010). GPS in the classroom: Using rubrics to increase student achievement. *Research in Higher Education Journal*, 9, 1-10.