



American Journal of Educational Research and Reviews (ISSN:2474-9265)



Rubrics 101: What, When and How

Larry Beard, Joesph Akpan, Charles E. Notar (Emeritus), Larry Beard

Jacksonville State University

ABSTRACT

Alternative assessment is any type of assessment other than standardized tests/ achievement tests. Alternative assessments include observations, interviews, record reviews, and performance reviews that are less structured than formal assessments and may not be validated or tested for reliability. Examples of alternative assessments include portfolios, interest inventories, work samples, journals, observations, checklists, teacher made tests, and anecdotal records. This article pertains to the scoring of these types of alternative assessment using rubrics. As the title of this article implies, creating rubrics, rules for writing, scoring systems (checklists, rating scales, and holistic) will be provided along with a comparison of the three and errors that can be found that will decrease the validity of findings.

Keywords: alternative assessment, checklists, holistic rating, rating scales, rubrics

*Correspondence to Author:

Dr. Charles E. Notar (Emeritus)
Jacksonville State University

How to cite this article:

Larry Beard, Joesph Akpan,
Charles E. Notar (Emeritus). Ru-
brics 101: What, When and How.
American Journal of Educational
Research and Reviews, 2019,4:44.

 **eSciPub**
eSciPub LLC, Houston, TX USA.
Website: <http://escipub.com/>

Introduction

A rubric is a scoring guide with criteria for evaluating students' work in direct relation to one or more of the program's learning outcomes and a rating scale indicating differing levels of performance. There are several types of rubrics including checklists, rating scales, holistic rubrics and descriptive or analytic rubrics. Checklists as per the term basically is a list of defined items that faculty may look for either in terms of grading an assignment or assessing a knowledge or skill embedded within the assignment. Similarly, rating scales also consists of a list of items but with a scale to indicate the degree to which the items indicated in the list are present. Holistic rubrics don't specify a list but rather have degree of performance or quality levels with brief narrative descriptions on what each performance level must contain. In other words, holistic rubrics assign a single score to an assignment or student work in which the criteria for the assignment is viewed as a whole.

Descriptive or analytic rubrics contain three components: a) a defined criteria, dimensions or components of an assignment, b) performance or quality level scale and c) a description box for each performance level and criteria. Analytic rubrics help identify strengths and weaknesses of students within the stated criteria/dimensions with analytic rubrics, each criterion is evaluated separately. By evaluating each criterion separately, the teacher receives specific diagnostic information, can give formative feedback to students, and provides an easier to link to instruction than holistic rubrics, are good for formative assessment, are adaptable for summative assessment and, if needed, provides an overall score for grading, as you can combine the scores.

However, analytic rubrics take more time to score than holistic rubrics and take more time to achieve inter-rater reliability than with holistic rubrics. Whereas with holistic rubrics, scoring is faster than with analytic rubrics, requires less time to achieve inter-rater reliability, and are

good for summative assessment. However, a single overall score does not communicate information about what to do to improve, so they should not be used for formative assessment (Brookhart & Nitko, 2008). The best rubrics for assessing student learning outcome are those that contain measurable criteria with descriptors that are well defined, objective and mutually-exclusive (Office of Institutional Research and Assessment, 2017; King, & Baxter Magolda, 2005; Mertler, 2001).

An instructional rubric is usually a one- or two-page document that describes varying *levels* of quality, from excellent to poor for a specific assignment. It is usually used with a relatively complex assignment, such as a long-term project, an essay, or a research paper. Its purposes are to give the students *informative feedback* about their works in progress and to give detailed evaluations of their final product (as cited in 1).

A rubric is an assessment tool that lists the *criteria* for a piece of work or counts and articulates gradations of quality for each criterion, from excellent to poor. Rubrics enable teachers to give more *informative feedback* to students. The *level* at which they met each criterion on the rubric can be checked with individual comments. Those same rubrics can be used as the bases of formative peer and self-assessments. Rubrics help to keep teachers fair and unbiased. The criteria are clearly stated. Ancillary factors, like how hard the student tried, are less likely to influence an instructor when the work is assessed with a rubric in hand (as cited in 2).

Developing rubrics for a variety of accomplishments

In general, performance tests require four types of accomplishments from learners:

Products: Poems, essays, charts, graphs, exhibits, drawings, maps, etc.

Complex cognitive processes: Skills in acquiring, organizing, and using information

Observable performance: Physical movements as in dance, gymnastics, or typing; oral presentations; use of specialized equipment as in focusing a microscope; following a set of procedures as when dissecting a frog, bisecting an angle, or following a recipe.

Habits of mind and mental and behavioral habits, group work, and social skills: recognition skills. As this list suggests, the effect of your teaching may be realized in a variety of ways (as cited in 3).

Creating Rubrics

Rubrics are becoming increasingly popular with educators moving toward more authentic, performance-based assessments. To make rubric more valuable, the rubric design process should engage students in the following steps:

1. Look at models: Show students' examples of good and not-so-good work. Identify the characteristics that make the good ones good and the bad ones bad. Identify and justify the best assessment method by which to gather evidence of proficiency on outcomes/behaviors in the domain. (A product? An observation of actual behaviors or demonstrations? An essay? A portfolio?)
2. List criteria: Use the discussion of models to begin a list of what counts in quality work. Specify whether an analytic or holistic scoring rubric will be used, or some combination of the two.
3. Articulate gradations of quality: Describe the best and worst levels of quality, then fill in the middle levels based on your knowledge of common problems and the discussion of not-so-good work. Specify the scoring scheme or point-allocation scale to be used. Checklist, rating scale, or both? Develop a draft of a scoring rubric to score the responses. Use observable descriptors to define the quality of a response at different levels of acceptability. Try to operationally define each scale point.
4. Practice on models: Have students use the rubrics to evaluate the models you gave them in Step 1. Identify weights in terms of point values

to be allocated to different components. This step is particularly relevant for analytic rubrics.

5. Use self- and peer-assessment: Give students their assignment. As they work, stop them occasionally for self- and peer-assessment. Design an exercise to capture the specified outcomes and behaviors in the domain. Write the instructions and prompts to elicit responses; set the context/conditions; provide instructions for assessors.

6. Revise: Always give students time to revise their work based on the feedback they get in

Step 5.

7. Use teacher assessment: Use the same rubric students used to assess their work yourself. Try out the task on a sample of typical respondents. (Or, alternatively, perform the task yourself) Observe or gather data on the possible range of responses to the task(s).

8. List common errors, emissions, or inaccuracies that you find in typical responses, based on an error analysis, revise the descriptors of scale points on your rubric. Use the actual responses to tighten and clarify observable indicators at different levels quality.

9. Attempt scoring actual responses using the revised rubric. Evaluate the extent to which the rubrics facilitate consistent scoring by different, but knowledgeable, raters.

10. Check back to make sure that rubric matches with indicators originally specified in the domain, as this will ensure content-based validity of the results.

11. Hold on to some sample answers at each score point to use as "anchor papers" during scoring. This will enhance reliability by controlling scorer errors (as cited in 4-12).

Rules for Writing Scoring Rubrics

1. Use a different rubric for each of the identified performance dimensions.
2. Limit the rubric to 3 to 6 points indicating levels of competence. Create only as many levels as you can describe given the complexity of the work.

3. Begin with clear performance criteria related to each dimension of performance and related to your learning objectives.
4. Write clear descriptions of the performance levels represented by each point. These are general descriptions of each level of quality. Some rubrics give lists of discrete characteristics for different levels of quality; usually, however, these are simply rating scales disguised as rubrics since students might do well on one criterion but poorly on another. Therefore, if performance criteria are listed at the top of rubrics and then performance levels are holistic quality descriptions, rubrics are easier to apply.
5. Make certain that each performance level is distinct from each other level. Create separate scoring rubrics for each dimension of performance you want to target. Students should be able to see how their work matches the score received and what needs to be done to perform at a higher level (as cited in 13-14).

Rules for Writing Scoring Rubrics

1. Use a different rubric for each of the identified performance dimensions.
2. Limit the rubric to 3 to 6 points indicating levels of competence. Create only as many levels as you can describe given the complexity of the work.
3. Begin with clear performance criteria related to each dimension of performance and related to your learning objectives.
4. Write clear descriptions of the performance levels represented by each point. These are general descriptions of each level of quality. Some rubrics give lists of discrete characteristics for different levels of quality. List performance criteria at the top of rubrics and then performance levels are holistic quality descriptions, rubrics are easier to apply.
5. Make certain that each performance level is distinct from each other level. Create separate scoring rubrics for each dimension of performance you want to target. Students should be able to see how their work matches the score received and what needs to be done to perform

at a higher level (as cited in Taylor, & Nolen, 2005; 31-37).

Choosing a Scoring System That Best Suits the Type of Accomplishment You Want to Measure

In general, there are three categories of rubrics to use when scoring performance tests: checklists, rating scales, and holistic scoring. Each has certain strengths and limitations, and each is more or less suitable for scoring products, cognitive processes, performances, and social skills.

Checklists

Checklists can be an efficient method of evaluating the level of skill mastery. Teachers can develop a checklist of skills, arranged in a consistent manner to systematically, quickly, and efficiently record whether specific skills or behaviors are or are not present. The purpose of the checklist should determine the kind of checklist that is used.

Checklists contain lists of behaviors, traits, or characteristics that can be scored as either present or absent. They are best suited for complex behaviors or performances which can be divided into a series of clearly defined, specific actions. Dissecting a frog, bisecting an angle, balancing a scale, making an audio tape recording, or tying a shoe are behaviors that require sequences of actions which can be clearly identified and listed on a checklist

Formats may vary.

Checklists can be used for an entire class or small groups, so that teachers can keep track of multiple students on one form rather than maintaining individual folders for each student being monitored. In this way, teachers can, at a glance, determine who does and does not need assistance in a specific area, a determination that is useful for instructional planning and program evaluation. In contrast, individual checklists are more useful for noting a particular student's strengths and weaknesses

To be most practical, checklists should be specific and have a realistic number of attainable

goals. The evaluative criteria should also be limited for ease of rating and scoring. Checklists are scored on a yes/no, present or absent, 0 or 1 points basis and should provide the opportunity for observers to indicate that they had no opportunity to observe the performance. Some checklists also include frequent mistakes that learners make when performing the task. In such cases, a Score of + 1 may be given for each positive behavior, -1 for each mistake, and 0 for no opportunity to observe. Other indicators may include evidence of completion (e.g., x = finished, 0 = not finished; yes or no), qualitative criteria (e.g., excellent, good, fair, poor), or relative level of proficiency (e.g., M = mastered, E = emerging, NS = no skill) (as cited in Arter & McTighe, 2001; 15-20).

Checklist Summary

List of key attributes of good performance checked present or absent.

Quick; useful with large number of criteria

Results can lack depth.

Rating Scales

Rating scales are typically used for those aspects of a complex performance which do not lend themselves to yes/no or present/absent type judgments. The most common form of a rating scale is one which assigns numbers to categories of performance. This scale focuses the rater's observations on certain aspects of the performance (accuracy, logic, organization, style, etc.) and assigns numbers to five degrees of performance (as cited in 21-22).

Rating Scales Summary

Performance continuum mapped on several point numerical scale ranging from low to high

Can record judgment and rationale with one rating.

Can demand extensive expensive development and training for raters.

Holistic scoring

Holistic scoring is used when the rater is more interested in estimating the overall quality of the performance and assigning a numerical value to

that quality than assigning points for the addition or omission of a specific aspect of performance. Holistic scoring is typically used in evaluating extended essays, term papers, or some artistic performances such as dance or musical creations (as cited in 23-24). Whereas holistic scoring is a faster method to assess students, holistic scoring can fail to establish weaknesses in student learning (Brookhart, 2013).

Combining Scoring Systems

Good performance tests require learners to demonstrate their achievements through a variety of primary traits, for example, cooperation, research, delivery, etc. Several ratings, therefore; may need to be combined from checklists, rating scales, and holistic impressions to arrive at a total assessment (as cited in 25).

Comparing the Three Scoring Systems

Each of the three scoring systems has its particular strengths and weaknesses. The following are considerations in developing a scoring system:

1. *Ease of construction.* Refers to the time involved in generating a comprehensive list of the important aspects or traits of successful and unsuccessful performance. Checklists, for example, are particularly time-consuming, while holistic scoring is not.
2. *Scoring efficiency.* Refers to the amount of time required to score various aspects of the performance and sum these scores into an overall score.
3. *Reliability.* The likelihood of two raters independently coming up with a similar score; or the likelihood of the same rater coming up with a similar score on two separate occasions.
4. *Defensibility.* Refers to the ease with which you can explain your score to a student or parent who challenges it.
5. *Quality of feedback.* Refers to the amount of information that the scoring system gives to learners or parents concerning strengths and weaknesses of their performance (as cited in 26-30).

Scoring

Limit the number of points that the assessment or component of the assessment is worth to that which can be reliably discriminated.

For example: 25 points assigned to a particular product or procedure assumes that the rater can discriminate 25 degrees of quality. When faced with more degrees of quality than can be detected a typical rater may assign some points arbitrarily reducing the reliability of the assessment. No results

Checklists

You create a list of specific steps in a procedure or specific behaviors. You check each behavior that occurs. The list may be long

These are primarily useful if the behaviors are in a sequence or if all the sub-tasks that make up the complete performance can be listed.

You create standards or criteria for evaluating a performance. Each standard has levels of competence, and you rate students according to how well they performed each standard as they complete the task

These are especially useful if each standard can be judged according to the level of the degree of quality rather than as simply being present or absent (as cited in 38-40).

Rules for Developing Checklists and Rating Scales

1. Make sure the checklist or rating scale has a clear focus, based on the performance criteria.
2. For checklists or rating scales for internal states (conceptual understanding, thinking skills, attitude, etc.), list the behaviors that will allow you to make the necessary inferences. Providing information on target behaviors will help students understand what they must do to demonstrate their competence. It is especially important if others (parents, aides, peers) will be using the rating scale.
3. For specific knowledge (e.g., knows letters of alphabet) or procedural knowledge (e.g., how to focus a microscope), list the specific knowledge

or procedural steps used to complete the task you plan to observe.

4. Use checklists for aspects of the performance that are all or nothing: Use terms like "yes/no," "present/absent," or "observed/not observed."

5 Use rating scales for performance aspects that can be judged in degrees. Use words to anchor the points of the rating scale relevant to the dimension of the performance to be assessed. Dimensions may relate to quality (e.g., effective, adequate, unacceptable), opinion (e.g., strongly agree, agree, disagree, strongly disagree), frequency (e.g., often, occasionally, rarely, never), degree of support (well supported, adequately supported, needs more support), accuracy (completely, mostly, partially), and others. Think carefully about what you want to know and how you want to look at it before deciding which dimension to use for your rating scale (as cited in Taylor & Nolen, 2005; 41-44).

Types of Rating Scales

Numerical Rating

Descriptive Rating

Descriptive Graphic Rating

Common Errors in Rating

Personal bias

Halo effect

Logical error

Personal Bias Error

Personal bias errors are indicated by a general tendency to rate all individuals at approximately the same position on the scale.

Generosity Error

Severity Error

Central Tendency Error

Central Tendency Error: When using a Likert scale that has an odd number you will be providing the rater with negative to the left and positive to the right and a "no rating" in the center. This "no rating" is the error of central tendency.

Use even numbers, which requires a “forced choice.” The rater must rate negative or positive because there is no middle ground rating.

Halo error occurs when a rater’s general impression of a student influences the rating of individual characteristics. This error can result in favorable or unfavorable ratings.

Logical error results when two characteristics are rated as more alike or less alike than they actually are because of the rater’s belief concerning their relationship.

All three types of errors can be markedly reduced by proper design and use (as cited in 45-48).

Rating Scale Errors and How to Avoid Them

You should be aware of several common errors that occur when teachers rate students. Teachers who do not use all of the points on a rating scale cause the following errors:

- Leniency error** occurs when a teacher tends to make almost all ratings toward the high end of the scale, avoiding the low end.
- Severity error** is the opposite of leniency error: A teacher tends to make almost all ratings toward the low end of the scale.
- Central tendency error** occurs when a teacher hesitates to use extremes and uses the middle part of the scale only. Central tendency errors sometimes occur when a teacher has to make strong inferences about a student (e.g., regarding “creativity” or “dedication”) and, in hesitation, the teacher tends to mark nearly everyone as average. Central tendency errors may occur when a teacher does not know the students very well.

Using only certain parts of the rating scale has two negative consequences. First, when you give only very high, very low, or “middle” ratings, you introduce your own quirks and biases into the ratings, thus lowering their validity for describing students’ ability in performing the task. Second, when your ratings bunch up and do not distinguish one students’ performance from another, they come unreliable, which in turn

reduces the validity of the scores (as cited in 49-53).

Here are other common teacher rating errors.

A **halo effect** occurs when a teacher lets her general impression of the student affect how she rates the student on specific dimensions. For example, if you gave a student a higher rating for his project than the student deserves because you “just know” that the student is “really” very good, you would be committing the halo effect error.

The general “halo” you place around the student affects your ability to judge the students standing on specific performances. (The halo effect may work in reverse, of course: Your general impression of a student as “not very good” may lead you to lower ratings on specific dimensions more than the student deserves.) One expression of the halo effect may occur when teachers need to make grading decisions for students whose assessment results put them on the border between two letter-grade categories: The error is that individuals who favorably impress a teacher are moved into the upper category; those who less favorably impress the teacher are moved into the lower category.

Personal bias occurs when a teacher tends to rate based on inappropriate or irrelevant stereotypes favoring boys over girls, whites over blacks, working families over welfare recipients, or particular families and individual students a teacher may dislike.

A **logical error** occurs when a teacher gives similar ratings on two or more dimensions of performance that the teacher believes are logically related but that are in fact unrelated. For example, a teacher may falsely believe that students with exceptionally high scores on scholastic aptitude tests also should be the top students in all subject areas. The teacher then marks the high-scoring aptitude test students differently from the way the low scorers are marked.

Logical errors are a result of a teacher's ignorance and unfounded beliefs, rather than

the teacher's personal quirks and biases about individuals or groups of students (as cited in 54-56).

Other errors occur when "outsiders" rate performance assessments. When states and large school districts implement performance assessments, individuals other than their teachers usually rate students' work. In these cases, the raters are trained in and practice using a particular scoring rubric.

- **Rater drift** occurs when the raters, whose ratings originally agreed, begin to redefine the rubrics for themselves. As a result, the raters no longer produce ratings that agree, with the original rubrics even though they were trained on the 'same rubrics. The remedy for this is to monitor the ratings and to retrain those raters who appear to have drifted away from agreed-upon standards (57-58).

- **Reliability decay** is a related error: Immediately after training, raters apply the rubrics consistently across students and mark consistently with one another. However, as time passes, the ratings become less consistent, both across students and across raters. Monitoring and retraining are remedies for this effect, too (as cited in 59-60).

Conclusion

Checklists, rating scales and rubrics are tools that state specific criteria and allow teachers and students to gather information and to make judgements about what students know and can do in relation to the outcomes. They are used to assess children's learning **strengths and weaknesses** in curriculum objectives and document progress in learning (as cited in 61,62)

References

1. retrieved from www.umsl.edu/.../most-03-2012/using-rubrics-to-promote-learning.pdf
2. retrieved from www.cs.northwestern.edu/~riesbeck/rubrics-rant.pdf
3. retrieved from <https://es.scribd.com/document/266162100/assessment-for-learning>
4. retrieved from https://tccl.arcc.albany.edu/knilt/images/6/63/Andrade_Rubrics.pdf
5. retrieved from www.ucdenver.edu/.../Documents/Tutorials/Rubrics/3_creating/index.htm
6. retrieved from <https://www.reference.com/How-To/Create/Rubrics>
7. retrieved from www.csulb.edu > Academic Technology Services
8. retrieved from <https://www.brightubeducation.com/teaching-methods-tips/23197-a...>
9. retrieved from <https://www.reference.com/How-To/Create/Rubrics>
10. retrieved from <https://academics.lmu.edu/.../rubrics/creatingarubric>
11. retrieved from <https://ctl.yale.edu/Rubrics>
12. retrieved from <https://go.view.usg.edu/shared/Documentation/9.4.1/Instructor/9.4.1...>
13. retrieved from map.mathshell.org/download.php?fileid=1165
14. retrieved from www.volstate.edu/sites/default/files/documents/human-resources/...
15. retrieved from <https://dennissparks.wordpress.com/2013/04/25/the-power-and-uses...>
16. retrieved from www.csun.edu/~krowlands/Content/Academic_Resources/Assessment...
17. retrieved from <https://checklist.com/classroom-checklist>
18. retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2047304/>
19. retrieved from www.stmarysbellinghen.nsw.edu.au/file.php/1/Teaching_and_learning..
20. retrieved from <https://bizfluent.com/info-8199048-advantages-disadvantages-using...>
21. <https://quizlet.com/40670189/hr-chapter-8-flash-cards>
22. retrieved from www.umaryland.edu/media/umb/hrs/eelr/pdp_rating_scale_definition.doc
23. retrieved from <https://www.ets.org/Media/Research/pdf/RR-06-07.pdf>
24. retrieved from <https://evaluationcanada.ca/secure/11-2-061.pdf>
25. retrieved from tarheels7.blogspot.com
26. retrieved from www.learnalberta.ca/content/mewa/html/assessment/checklists.html
27. retrieved from hosted.jalt.org/pansig/2004/HTML/Nakamura.htm
28. retrieved from <https://www.scribd.com/presentation/182405616/classroom...>
29. retrieved from <https://pampots28.wordpress.com/2013/12/04/advantages-and...>

30. retrieved from <https://www.scribd.com/presentation/182405616/classroom...>
31. retrieved from <https://reliablerubrics.com/.../09/rules-for-writing-quality-criteria>
32. retrieved from <https://www.lessonplanet.com/article/literature/rules-for-rubrics>
33. retrieved from [https://www.eNow.com/ScoringRubrics For Writing](https://www.eNow.com/ScoringRubricsForWriting)
34. retrieved from <https://www.pareonline.net/getvn.asp?v=7&n=3>
35. retrieved from toeflpro.blogspot.com/2011/10/scoring-rubrics-what-you-need-to...
36. retrieved from education.ucdavis.edu/.../march_152014_am_rubrics_additional.pdf
37. retrieved from <https://www.writetolearn.net/scoring-rubrics>
38. retrieved from <https://blog.performancetask.com/how-will-we-evaluate-student...>
39. retrieved from <https://hr.berkeley.edu/hr-network/central-guide-managing-hr..>
40. retrieved from <https://www.opm.gov/.../planning/developing-performance-standards>
41. retrieved from <https://apps.support.sap.com/sap/support/knowledge/public/en/...>
42. retrieved from <https://peoplecentre.wordpress.com/2013/05/13/designing...>
43. retrieved from <https://www.coursehero.com/file/p431dfe/Identify-and-describe-a-variety-of-performance-rating-scales-that-can-be-used/>
44. retrieved from www.ncbi.nlm.nih.gov › Journal List › Front Psychol › v.4; 2013
45. retrieved from <https://inkessays.com/21149-2>
46. retrieved from <https://www.slideshare.net/ciplashine/rating-scales-31378095>
47. retrieved from methods.sagepub.com/reference/encyclopedia-of-survey-research...
48. retrieved from www.dartmouth.edu › ... › Performance Management
49. retrieved from <https://interviewstar.wordpress.com/2011/01/04/leniency-strictness...>
50. retrieved from <https://thelawdictionary.org/central-tendency-error>
51. retrieved from www.coursehero.com › MGMT
52. retrieved from <https://psychologydictionary.org/leniency-error>
53. retrieved from <https://www.tuturself.com/posts/view?menuId=116&postId=933>
54. retrieved from <https://smartchurchmanagement.com/performance-appraisal-rater-errors>
55. retrieved from <https://www.happyoffice.io/blog/biases-in-performance-appraisals...>
56. retrieved from learnthat.com/5-types-of-bias-in-a-performance-review
57. retrieved from <https://www.merriam-webster.com/dictionary/rater>
58. retrieved from <https://www.brentwoodindustries.com/.../what-does-drift-rate-mean>
59. retrieved from images.pearsonassessments.com/images/tmrs/P-performance_Scoring_for...
60. retrieved from ii.library.jhu.edu/.../sharing-assignment-rubrics-with-your-students
61. retrieved from <https://oira.unc.edu/files/2017/07/Developing-and-Using-Rubrics.pdf>
62. retrieved from <https://www.york.cuny.edu/president/institutional-effectiveness/institutional-assessment-1/academic-assessment/program-assessment-tools/rubrics>
63. Brookhart, S. & Nitko, A. (2008). *Assessment and Grading in Classrooms*. Upper Saddle River, NJ: Pearson Education.
64. Brookhart, S. (2013) *How to Create and Use Rubrics for Formative Assessment and Grading*. Alexandria, VA: ASCD
65. King, P.M. & Baxter Magolda, M.B. (2005). A developmental model of intercultural maturity, *Journal of College Student Development*, 46(2), 571-592.
66. Mertler, Craig A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Available online: <http://www.pareonline.net/getvn.asp?v=7&n=25>
67. Office of Institutional Research and Assessment. (2017). *Using rubrics to assess student learning outcomes at the program level*.
68. Taylor, C. S., & Nolen, S. B. (2005). *Classroom Assessment: Supporting Teaching and learning in Real Classrooms*. Columbus, OH: Pearson/Merrill Prentice Hall, p. 101, 241.