Using an Assessment Rubric to Evaluate an Online Learner's Performance

Kimberly Payne
Imedia.it
Houston, Texas
kim.payne@imediait.com

ABSTRACT

Contractors were previously limited in after action reviews (AAR), due to certain statement of work regulations, to offering the learners generalized feedback related to performance in Web-based training (WBT). Not directly addressing learner performance through evaluation responses made it difficult to determine the learner's strengths and weaknesses. One possible solution is to use a detailed assessment rubric. An assessment rubric is a form of evaluation that graphically depicts the conditions and standards for a training product. A learner knows exactly what is expected of him or her. In order for the assessment rubric to benefit distance education participants, the rubric should be presented at the beginning of the training and in the AAR. Allowing the learner to review the conditions and standards of the training at the beginning of the event informs the learner of learning expectations. Using the assessment rubric as the AAR, with the learner's results clearly articulated, allows the learner to analyze his or her results. Presenting the assessment rubric graphically helps the learner determine the skills in which he or she is proficient and which skills may need some more development. To design an assessment rubric, the first step is to identify the learning objectives that the learner must successfully meet. The second step is to determine the levels of performance and the criteria for each level. Once the learning objectives, performance levels, and criteria have been determined, the assessment rubric needs to be reviewed by subject matter experts and education specialists. If an assessment rubric is created before the training product, the rubric should be reviewed again so as to ensure its accuracy and appropriateness. An assessment rubric can be a valuable tool for evaluating learner performance if it is well-planned, detailed, and effectively integrated into WBT.

ABOUT THE AUTHOR

Kimberly Payne is an Instructional Designer for Imedia.it. Prior to her job at Imedia.it, she was a professional classroom educator in the Alvin, Texas Independent School District. Kim received her Master's degree in Instructional Technology from the University of Houston – Clear Lake. Since joining the Imedia.it team in 2001, she has worked on both corporate and military projects. Kim was the lead instructional designer for the 97B10 Web-based training. She has also developed several MASINT courses for INSCOM. Mrs. Payne recently completed the Installation Force Protection (IFP) Project for the U.S. Army Intelligence Center and School located at Fort Huachuca. Her current projects include the S2 course and Intelligence in Combating Terrorism. Mrs. Payne has presented at the Region IV Technology Applications Workshop, Training and Learning Week 2003, and the Distance Education Conference sponsored by Texas A&M's Center for Distance Learning Research. Her chapter "Assessment Elements in Web-based Training" has been accepted for publication in *Online Assessment and Measurement, Volume II: Case studies from higher education, K-12, and corporate.*

Using an Assessment Rubric to Evaluate an Online Learner's Performance

Kimberly Payne
Imedia.it
Houston, Texas
kim.payne@imediait.com

INTRODUCTION

As technology rapidly advances, so does the training medium. Before the advent of the Information Age, distance learning merely consisted of simple text on a Web page. The content was generally followed by some type of knowledge-level assessment. These assessments consisted of multiple-choice, true/false, and matching questions. Such assessment methods have advanced with the addition of audio and other media forms. Now training can be scenario-based and offer learners multiple ways to experience the courseware. Assessments can measure the performance of the learner without being limited to traditional knowledge-level exam questions. Once the assessment types shifted, the feedback to the learner and the review of their work had to change. Contractors were limited in most after-action reviews (AAR), due to certain statement of work (SOW) regulations, to offering the learner generalized feedback related to performance in web-based training (WBT). Not directly addressing student performance through evaluation responses made it difficult to determine the learner's strengths and weaknesses. How can contractors give learners the detailed feedback that they need in an online setting? One possible solution is to use a detailed assessment rubric. In this paper, we will examine what is defined as an assessment rubric, how it is constructed, and how to appropriately integrate it into WBT.

WHAT IS AN ASSESSMENT RUBRIC?

An assessment rubric is a form of evaluation that graphically depicts the conditions and standards for a training product. As stated by Pickett and Dodge (2001), "The rubric is an authentic assessment tool which is particularly useful in assessing criteria which are complex and subjective. Authentic assessment is geared toward assessment methods which correspond as closely as possible to real world experience." However, a rubric is not the appropriate tool for every assessment situation. Rubrics are meant to be used toward the evaluation of portfolios, essays, multimedia presentations, and comparable deliverables; "where and when a scoring rubric is used depends . . . on the purpose of the assessment" (Moskal, 2000). When an assessment rubric is the most appropriate evaluative solution, the primary concern is to "address the aspects of student work that you feel are most important" (Chicago Public Schools, 2000).

Pickett and Dodge state that "The rubric is one authentic assessment tool which is designed to simulate real life activity where learners are engaged in solving real-life problems" (2001). One beneficial aspect of an assessment rubric is the level and quality of thinking reflected in the student product. Most educators are very familiar with Bloom's Taxonomy. In Bloom's model, there are six levels of learning:

- Knowledge (lowest level)
- Comprehension
- Application
- Analysis
- Synthesis
- Evaluation (highest level) (Language and Learning Improvement Branch)

The knowledge level is the most basic and only requires students to define or recite knowledge. At the comprehension level, the learner must be able to discuss, describe, or explain pieces of knowledge. A learner at the application level must be able to solve problems and apply knowledge within a real-life situation. Learners at the analysis level would have to be able to group and arrange knowledge. Synthesis is the fifth level and requires that a learner can predict and develop knowledge. The highest level is evaluation, which requires that a learner be able to judge and evaluate knowledge. As might be presumed, it is difficult, and often inappropriate, for an instructor to consistently offer instruction at the highest level of Bloom's Taxonomy. Yet when the instructor focuses upon the higher levels of knowledge understanding and manipulation in Bloom's Taxonomy, appropriate methods of evaluation must be integrated. This has been a problem in the American school system, in which "many of the skills we want our schools to impart to young people are not well measured by traditional multiple choice tests" (Wise, 1993, p. 1).

An assessment rubric is "designed more to assess proficiency at practiced skills than to measure simpler factual knowledge" (Wise, 1993, p. 1).

Crawford states that "A rubric guides the expectations for the assignment" (2001). As such, the focus of the assessment rubric is to ensure that the learner knows exactly what is expected of him or her. The learner is presented with the information and allowed to decide how much effort to put forth. Pickett and Dodge state that "Rubrics clearly show the student how their work will be evaluated and what is expected" (2001) although some educators might argue that this approach stifles creativity. It can stifle creativity if the assessment rubric is too detail-oriented and lessens the creativity of the learner; however, an appropriate and successful assessment rubric offers an appropriate balance between specific and general guideline expectations. An assessment rubric is not meant to delineate assignment expectations ad nauseum; it is, however, meant to present and delineate the criteria through which the assignment will be evaluated. For example, an instructor assigns a writing assignment and allows the learner to choose the topic, so as to fit within the learner's area of comfort and expertise. The instructor utilizes an assessment rubric that is presented and explained to the learners. The assessment rubric may address only issues related to grammar, punctuation, and writing style; in this situation, the learner has complete reign over content and can be creative. The instructor is looking more towards style and knowledge about language rules and applications. Another example would encompass an instructor assigning a research paper assignment, which also allows the student to define the parameters of the assignment (e.g., topic, breadth of subject, length). The instructor also utilizes an assessment rubric that is presented and explained to the learners; however, significantly different from the previous example, the assessment rubric offers significantly more detail (e.g., referential structure, number of referential quotes, manuscript format, spelling grammar, punctuation, author voice, writing style, transition structure). Although both example, offer similar writing projects, significantly different delineations within the assessment rubric enhance the assignment and denote differing levels within Bloom's Taxonomy (e.g., application versus evaluation levels).

Advantages

There are many advantages to using an appropriate assessment rubric. First, "rubrics make the instructors clarify his or her criteria in specific terms" (Pickett & Dodge, 2001). By clarifying the criteria, an instructor is more consistent when grading learner work. In WBT, the learner's performance can be tracked through a database format. The database can record which scenarios the learner has completed, how he or she scored on each assessment, and anything else that is relevant to a learner's performance. This information can then be delineated by performance levels outlined in the assessment rubric. This is not the case when traditionally evaluating a learner's understanding of the information, due to the lack of higher order thinking skills encompassed within traditional evaluative methods. For example, an instructor may assign a compare and contrast paper as a cumulative activity. Once learners have turned in the paper the instructor begins evaluating the work, difficulties may arise; without an assessment rubric laying out the exact grading criteria, the instructor may have a hard time consistently evaluating the learner products.

A second advantage is that a "rubric allows assessment to be more objective and consistent" between instructors (Pickett & Dodge, 2001). For example, in a live interview situation, the instructors are the responsible parties for role playing a scenario and taking notes related to the learner's performance. Because the instructors are focused upon numerous components at the same time, they may overlook important aspects of performance and there is no simple way to accurately evaluate every aspect of a learner's performance. Instructors have different experiences and different underlying educational philosophies; as such what one instructor may evaluate as being of primary importance, another instructor may overlook. Learners are at a disadvantage when there is a lack of consistent feedback. Utilizing detailed tracking of a learner's performance, as outlined in the performance criteria of the assessment rubric, is a significant advantage and greatly reduces learner feedback inconsistencies. When several instructors can implement an assessment rubric and offer learners approximately the same evaluative outcome on an assessment component, then the assessment rubric is considered reliable, valid, and consistent. "Ideally, two raters using the same rubric would agree at least 65 percent of the time, and be within one point of each other a full 98 percent of the time" (Boston, 2002, p. 25).

A third advantage is that "rubrics provide useful feedback regarding the effectiveness of the instruction" (Pickett & Dodge, 2001). Appropriate feedback "helps students maintain a sense of control, reduces uncertainty, and encourages a higher level of thinking" (Gregory & Kuzmich, 2004, p. 17). As such, useful feedback must be descriptive, so as to offer the learner detailed, constructive information. For example,

"Learner's work contains no grammatical errors," is preferable to "Learner's work is good." The terms being used should be quantifiable and not merely judgment calls. As well, the focus of the assessment should be upon positive aspects of the learner's product. "Rubrics can encourage assessing individual strengths versus viewing weaknesses…" (Taggart, Phifer, Nixon, Wood, 2001, p. 45). A detailed assessment rubric provides feedback to the learner about his or her current performance level and what things he or she could do to accomplish an advanced level of performance. This is especially important for WBT. In a traditional classroom setting, an instructor is present and may offer timely feedback. Learners engaged in online classes rarely, if ever, see an instructor. Learners who take online classes need detailed, immediate feedback related to their performance; just as if the learner were sitting in a classroom. A detailed assessment rubric is one way to provide this type of timely, detailed directive feedback to the learner.

Common Features

All assessment rubrics are different. Rubrics focus upon significantly different learning objectives, as well as audiences. "Rubrics offer the learner and the instructor the ability to review the expectations of any assignment so that everyone has a clear understanding of the final product, as well as the expected grading criteria that will be in place for each assignment" (Crawford, 2001). According to Pickett and Dodge (2001) all rubrics should contain the following features:

- Focus on measuring a stated objective (performance, behavior, or quality);
- Use a range to rate performance; and,
- Contain specific performance characteristics arranged in levels indicating the degree to which the standard has been met.

No matter what content is an integral component within the learning objectives, whether it is being presented over the computer or in a traditional classroom setting, all assessment rubrics should contain these three common features.

Sample Assessment Rubric: Installation Force Protection Contract

The assessment rubric that is described in the following sections was developed as part of the requirements associated with the Installation Force Protection (IFP) WBT. The IFP project is a course funded by Distance Learning Center at Fort Huachuca. The learning objective of the training is for the Garrison S-2 or Armory Commander to identify potential threats to an installation and recommend protective measures to prevent casualties and/or property damage. In order to accomplish this goal, the learner is required to successfully perform the simulated daily duties of an S-2 or Armory Commander. These learning objectives include:

- Determining geographical boundaries, where the friendly force is vulnerable to threat activity, battlefield environment, the threat and its capabilities, enemy course of action, and intelligence gaps;
- Evaluating matrices, the essential elements of friendly information, the intelligence, surveillance, and reconnaissance plan, the public affairs statement, and the intelligence summary;
- Revising the priority intelligence requirements, the operations and operational security annex plans; and,
- Implementing the plan against an attack.

The learning objectives of the course require the learner to function in the upper three levels of Bloom's Taxonomy (e.g., analysis, synthesis, evaluation). Based on the learning objectives and level of performance required by the learner, the rubric is the perfect solution through which to assess the learner's grasp of information and ability to meet the course learning objectives.

HOW IS AN ASSESSMENT RUBRIC DESIGNED?

There are many resources available that explain how to design an assessment rubric. Designing an assessment rubric is not a simplistic, nor swift, task. Many of the "how-to" sources suggest that an instructor look for an assessment rubric that has already been designed and adapt or modify it to fit the needs of the learning objectives. This is especially true for professional classroom educators who may have developmental time constraints. When an instructor decides to design an assessment rubric, these tips from the Chicago Public Schools (2000) should be kept in mind:

- Decide whether the rubric addresses the most important aspects of student performance;
- Decide whether or not the rubric addresses the instructional outcome(s) to be measured;
- Decide whether the rubric includes anything extraneous. If so, change the rubric or use a different one; and,
- Make sure the rubric is clear.

When designing an assessment rubric, the first step is to identify the objectives that the learner must meet so as to be deemed successfully trained, as displayed in

Table 1. This step is accomplished early in the development process. The content for the course must be delineated before an assessment rubric can be addressed. The terminal learning objective for IFP is located in the statement of work (SOW). Based on the terminal learning objective, the subject matter expert (SME) determined what content would be the focus of the training. When the content was delineated, the instructional designer and SME wrote the enabling learning objectives.

Table 1. Identify the learning objectives.

Task			
Determine geographical boundaries.			
Determine where the friendly force is vulnerable to threat activity.			
Determine the battlefield environment.			

The second step is to determine the levels of performance, as displayed in Table 2. Rubric designs vary by the subject matter and the audience, which are dependent upon the needs of the learning objectives and the components within the final product deliverable. In the IFP WBT, the assessment rubric comprises three levels of performance. "If greater distinctions are desired, then comparisons can be made between the criteria for each existing score level" (Moskal, 2000). Simkins suggests that four is the most appropriate number of evaluative levels (1999), because "The problem with using three levels is that you probably won't be able to make fine enough discriminations. If you try to force student work into five or more levels, you may find yourself splitting hairs" (Simkins, 1999). However, three levels of performance may be appropriate if engaged in training that focuses upon reaching specific levels of performance expertise.

Table 2. Determine the levels of performance.

Task	Below Performance Standard	Meets Performance Standard	Displayed Superior Performance
Determine geographical boundaries.			
Determine where the friendly force is vulnerable to threat activity.			
Determine the BFE.			

Step three is to determine the criteria for each level of performance, so as to "Identify the qualities that need to be displayed in a student's work to demonstrate proficient performance" (Moskal, 2000). It is usually easier to initially focus upon the top level of performance, as displayed in Table 3. A learner would be required to perform at the stated criteria level perfectly in order to achieve the displayed superior performance ranking. Remember that "good rubrics are neither too specific nor overly general" (Simkins, 1999).

Table 3. Determine the upper level of performance.

Task	Below Performance Standard	Meets Performance Standard	Displayed Superior Performance
Determine geographical boundaries.			Correctly determined and depicted the limits of the AO and AI on a map.
Determine where the friendly force is vulnerable to threat activity.			Correctly analyzed and prioritized the critical infrastructure and assets vulnerable to an enemy threat.
Determine the battlefield environment.			Correctly analyzed and evaluated the terrain, weather, and other battlefield characteristics to determine the effects on operations.

Once this has been successfully delineated, the next area of focus is upon the lowest indicator level of performance. "After defining the criteria for the top level of performance, the evaluator's attention may be turned to defining the criteria for the lowest level of performance" (Moskal, 2000). Following the advice of Moskal (2000), the IFP design team found it is easiest to make decisions about the lowest level next, as displayed in Table 4. This performance level is the lowest and requires very little effort on the part of the learner. If the learner chooses to perform on this level, it is considered significantly below performance standard (e.g., unsuccessful, a failure).

Table 4. Determine the lowest level of performance.

Task	Below Performance Standard	Meets Performance Standard	Displayed Superior Performance
Determine geographical boundaries.	Determined and depicted secondary geographical areas.		Correctly determined and depicted the limits of the AO and AI on a map.
Determine where the friendly force is vulnerable to threat activity.	Determined assets not critical to accomplishing the mission. Uncertain of the critical infrastructure of the garrison.		Correctly analyzed and prioritized the critical infrastructure and assets vulnerable to an enemy threat.
Determine the battlefield environment.	Limited knowledge of the terrain, weather, and other battlefield characteristics and what effects they have on operations.		Correctly analyzed and evaluated the terrain, weather, and other battlefield characteristics to determine the effects on operations.

Lastly, the focus is upon defining the criteria for the middle level of performance, as displayed in Table 5. Because the high and low levels have already been defined, the task is less strenuous, because "The contrast between the criteria for top level performance and bottom level performance is likely to suggest appropriate criteria for middle level performance" (Moskal, 2000).

Table 5. Determine the middle level of performance.

Task	Below Performance Standard	Meets Performance Standard	Displayed Superior Performance
Determine geographical boundaries.	Determined and depicted secondary geographical areas.	With additional training, defined and depicted the limits of the AO and AI on a map.	Correctly determined and depicted the limits of the AO and AI on a map.
Determine where the friendly force is vulnerable to threat activity.	Determined assets not critical to accomplishing the mission. Uncertain of the critical infrastructure of the garrison.	Determined a majority of the critical infrastructure and assets vulnerable to an enemy threat.	Correctly analyzed and prioritized the critical infrastructure and assets vulnerable to an enemy threat.
Determine the battlefield environment.	Limited knowledge of the terrain, weather, and other battlefield characteristics and what effects they have on operations.	Displayed general knowledge of the terrain, weather, and other battlefield characteristics and what effects they have on operations.	Correctly analyzed and evaluated the terrain, weather, and other battlefield characteristics to determine the effects on operations.

Once the learning objectives, performance levels, and criteria have been determined, it is appropriate that the assessment rubric be reviewed by the subject matter expert (SME) and education specialists. This step is important because the team must "agree on the score that should be assigned to a given piece of student work" (Chicago Public Schools, 2000), as related to concerns of reliability and validity.

HOW CAN AN ASSESSMENT RUBRIC BE INTEGRATED INTO WEB-BASED TRAINING?

In order for the assessment rubric to benefit distance education participants, the rubric should be "presented to the learner at the outset of the assignment so that both the learner and the instructor know exactly what is expected of them" (Crawford, 2001). Allowing the learner to review the conditions and standards of the training at the beginning of the training informs the learner of the training expectations. To accomplish this goal in IFP, the design team decided to introduce the learner to the assessment rubric during the tutorial section of the training. "The availability of the rubric to the learner at the beginning of the assignment aids the learner in refining and orienting the product to include the desired content aspects…" (Crawford, 2001). During the training, the learner may access the tutorial at any time, and so may view the assessment rubric.

An after-action review (AAR) is a report at the end of a training session. As was stated earlier, an AAR should be general and not give the learners the answers to the training, as per Army guidance. Table 6 displays a sample of an AAR taken from an Army course in which the learner had to conduct an interview. The left side of the report outlines every question that the learner asks; the right side of the report contains the feedback that the learner is given about the question. The problem with this AAR is that it does not offer the learner enough qualitative feedback related to their performance. The database tracked the questions that the learner asked and provided feedback about those questions. The learner is simply told whether or not the question is critical, good, or needs more work.

Table 6. Sample AAR from 97B10 WBT

Questions:		**AAR:**
INCIDENT:		
What did you order at the restaurant?		Good question.
How did you pay for your bill?		This question did not have relevance to the case.
Did you order an alcoholic beverage while dining?		Good question.
What did you see?		This question was not specific enough.
What did you see at the restaurant?		This was a critical question that yielded vital information to the case.
Who was Chief MANLEY with?		This was a critical question that yielded vital information to the case.

The problem is that the learner is left wondering: what questions should I have asked, what grade did I receive, and in what ways could I have done better? If this AAR had been designed as an assessment rubric, the learner would know the answers to these questions. Using an assessment rubric as the AAR for IFP, with the learner's results clearly marked as displayed in Table 7, allows the learner to thoroughly analyze his or her results. Presenting the performance feedback in a graphic format helps the learners determine exactly what skills they are proficient in and which skills may need some more work. Feedback should be offered to a learner throughout the training course, as well as at the end of the training to recap and review the learner's progress.

CONCLUSIONS

Assessment rubrics are an evaluative tool in which the "Benefits far outweigh the work of creating rubrics, and programs become stronger as a result of careful examination" (Taggart, et al. 2001, p. 27). An assessment rubric can be a valuable tool for evaluating learner performance if it is well-planned, detailed, and effectively integrated into WBT. This IFP WBT is in the production stage and should be completed in early summer, 2004. Validation results that will indicate the effectiveness level of the rubric as an authentic assessment tool. The requirement for an assessment rubric first appeared in the IFP SOW. Since then, the assessment rubric requirement has been displayed in several more statements of work, including the second half of the IFP project slated to start mid-summer, 2004.

Table 7. Sample Assessment Rubric as the AAR

Task	**Below Performance Standard**	**Meets Performance Standard**	**Displayed Superior Performance**
Determine geographical boundaries.	Determined and depicted secondary geographical areas.	With additional training, defined and depicted the limits of the AO and AI on a map.	Correctly determined and depicted the limits of the AO and AI on a map.
Determine where the friendly force is vulnerable to threat activity.	Determined assets not critical to accomplishing the mission. Uncertain of the critical infrastructure of the garrison.	Determined a majority of the critical infrastructure and assets vulnerable to an enemy threat.	Correctly analyzed and prioritized the critical infrastructure and assets vulnerable to an enemy threat.
Determine the battlefield environment.	Limited knowledge of the terrain, weather, and other battlefield characteristics and what effects they have on operations.	Displayed general knowledge of the terrain, weather, and other battlefield characteristics and what effects they have on operations.	Correctly analyzed and evaluated the terrain, weather, and other battlefield characteristics to determine the effects on operations.

ACKNOWLEDGEMENTS

I would like to acknowledge Luciano Iorizzo, Deputy Director of Army Training Support and Command, for having the foresight to envision an assessment rubric as part of the Installation Force Protection Training.

Thank you Helen Remily and John McGovern of the Learning Technology Branch at Fort Huachuca for supporting the development of the assessment rubric.

I sincerely appreciate all of Caroline Crawford's, professor at the University of Houston—Clear Lake, assistance with this paper.

REFERENCES

Boston, C. (2002). *Understanding Scoring Rubrics: A Guide for Teachers.* Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 471518).

Chicago Public Schools. (2000). *Ideas and rubrics.* Retrieved June 7, 2004, from http://intranet.cps.k12.il.us/Assessments/Ideas_and_Rubrics/Rubric_Bank/Other_Rubrics/other_rubrics.html

Crawford, C. (2001). *Rubrics: Models of evaluation within a constructivist learning environment.* Paper presented at the spring conference of the National Council of Teachers of English, Birmingham, AL.

Gregory, G. H., & Kuzmich, L. (2004). *Data driven differentiation in the standards-based classroom.* Thousand Oaks, CA: Corwin Press, Inc.

Language and Learning Improvement Branch – Division of Instruction. Use Bloom's Taxonomy to Plan Lessons and Questions [Brochure]. Maryland State Department of Education.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation,* 7(3). Retrieved June 7, 2004 from http://PAREonline.net/getvn.asp?v=7&n=3

Pickett, N., & Dodge, B. (2001). *Rubrics for Web lessons.* Retrieved June 7, 2004, from http://webquest.sdsu.edu/rubrics/weblessons.htm

Simkins, M. (1999). *Designing great rubrics.* Technology and Learning, 20 (1), 23-24, 28-30. Retrieved June 8, 2004, from http://www.highbeam.com/library/doc1.asp?DOCID=1G1:56457742&num=2&ctrlInfo=Round3a%3AProd%3ASR%3AResult

Taggart, G. L., Phifer, S. J., Nixon, J. A., & Wood, M. (Eds.). (2001). *Rubrics: A Handbook for Construction and Use.* Lanham, Maryland: ScarecrowEducation.

Wise, L. (1993). *Scoring rubrics for performance tests: Lessons learned from job performance assessment in the military.* Atlanta, GA. (ERIC Document Reproduction Service No. ED 358114).