Rubrics as Assessment, Evaluation and Scoring Tools
Azliza Muhammad, Othman Lebar & Siti Eshah Mokshein

To Link this Article:  http://dx.doi.org/10.6007/IJARBSS/v8-i10/5309  DOI: 10.6007/IJARBSS/v8-i10/5309

Received: 09 Sept 2018, Revised: 22 Oct 2018, Accepted: 29 Oct 2018

Published Online: 05 Nov 2018

In-Text Citation: (Muhammad, Lebar, & Mokshein, 2018)


Copyright: © 2018 The Author(s)
Published by Human Resource Management Academic Research Society (www.hrmars.com)
This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at: http://creativecommons.org/ licences/by/4.0/legalcode

http://hrmars.com/index.php/pages/detail/IJARBSS

Full Terms & Conditions of access and use can be found at
http://hrmars.com/index.php/pages/detail/publication-ethics

1417
Rubrics as Assessment, Evaluation and Scoring Tools

Azliza Muhammad, Othman Lebar & Siti Eshah Mokshein
Faculty of Human Development, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia
Email: muhammadazliza@gmail.com

Abstract
Various types of assessment have grown rapidly, in line with the development of the education sector. An assessment can be conducted using a project-based assessment, authentic assessment, creative assessment and technology-based assessment. Meanwhile, assessment tools consist of questionnaires, test papers, quizzes, scoring guides, and rubrics. Nowadays, people are turning their attention to using the rubric as an assessment tool where the details will be further discussed in this study, which includes the concepts of rubrics, rubric types, components of a rubric, rubric modeling and also assessment and scoring using rubrics.

Keywords: Rubric, E-Rubric, Assessment, Scoring Tools

Introduction
A rubric is an assessment tool that has a description of the expected performance for each criterion in order to achieve a grade or certain outcomes. Rubric is a systematic method to collect data regarding knowledge and skills as stated by Churches (2015) in his study. Garfolo (2016) agreed that rubrics can be used to measure certain behaviour. In detail, the rubric is a scale rating questionnaire with selected-response items (Haladyna & Rogriguez, 2013). The specific or standard expectations of a performance to evaluate learning outcomes (Aiken, 1996; Company et al., 2017; Stevens & Levi, 2013) are key part of the rubric as it does not only serve as a tool of assessment but also serves as a learning tool as quoted by Andrade and Du (2005).

“Rubrics can teach as well as evaluate”
Therefore, this obvious rubric application can benefit any discipline (Montgomery, 2002). However, as highlighted by Company and his colleagues (2017), the overly abstract-constructed rubric will make it difficult for lecturers and students to understand the criteria. Meanwhile, e-rubric which refers to an electronic rubric or a computer-assisted rubric is a rubric that is plugged-in to an electronic platform or an electronic rubric (Anglin, Anglin, Schumann, & Kaliski, 2008). As stated by Company and colleagues (2017), e-rubric serves as a formative assessment.
Benefits of Rubric
The rubric is beneficial in terms of:

- Saving time, facilitating and speeding up the feedback process (Churches, 2015; Jonsson & Svingby, 2007; Reynolds-Keefer, 2010; Stevens & Levi, 2013)
- Helping to explain learning goals and standards (Andrade & Du, 2005; Hattie & Timperley, 2007)
- Improving performance to achieve a set of standard (Moskal, 2000; Sadler, 2009)
- Allowing supervision and monitoring of student progress (Reddy & Andrade, 2010)
- Helping students to focus on their learning efforts, producing better quality work and assignments to achieve better grades (Andrade & Du, 2005)
- Providing more accurate and fair assessments, evaluations and grades are more transparent which can avoid personal prejudice (Andrade & Du, 2005; Ellis & Kelder, 2012; Isbell & Goomas, 2014; Rivas, De La Serna, & Martinez-Figueira, 2014)
- Their capacity as a form of communication (Andrade & Du, 2005)
- Response tool which focuses on ongoing tasks and can be used to grade final products (Andrade & Du, 2005; Rivas et al., 2014)
- Reducing worry or concern about assignments (Andrade & Du, 2005; Reynolds-Keefer, 2010) because the descriptions in the rubric can be matched with the students' abilities and maturity, which is easy to understand, and the feedback is more transparent and constructive (Churches, 2015).

The subjective nature of the rubric can be further reduced, thus enhancing the understanding of a rubric. This can be achieved by providing training on how to use the rubric (Lovorn, Michael & Rezaei, 2011).

Types of Rubrics
The nature of the assignment and the purpose of using rubric as a tool will determine which rubric is the most suitable one to use (Riddle & Smith, 2008). Rubric serves best as a tool for formative and summative assessments (Educational Reserach Service, 2004).

Analytic rubric and holistic rubric are the two most commonly featured rubrics. Researchers namely Haladyna and Rodriguez (2013) and Moskal (2000) further categorized the rubric into either task-specific or generic models. Rubric that is specific for an assignment is referring to the task-specific model. Meanwhile, a rubric that is used for a non-specific assignment but has similar criteria to a specific assignment is called a generic model. Haladyna and Rodriguez suggested that task-specific model is best used for classroom learning while the generic model is more suitable to be used for tests and examinations.

An analytic rubric is a set of performance scores for a number of different evaluation criteria that is independent of each other (Baryla, Shelley, & Trainor, 2012; Haladyna & Rodriguez, 2013). For analytical rubric, Haladyna and Rodriguez (2013) stated that it takes quite a time to develop the rubric, but the score will be more accurate. In order to help students and lecturers identify their performance levels through discrete criteria, analytic rubric is more appropriate to be used for signature assessment (Garfolo et al., 2016).

On the other hand, the holistic rubric combines all analytical features into one single score (Haladyna & Rodriguez, 2013). The holistic rubric is easier to develop but they do not
detail the level of student’s performance because only a single score is given to evaluate the student’s work (Baryla et al., 2012).

A checklist usually has a review section to justify whether an assignment is complete or not, which is similar to how a rubric works. As pointed out by Riddle and Smith (2008), in addition to analytic and holistic rubrics, a checklist is also named as a rubric.

Rubric Components
Typically, there are four components arranged in a rubric which are:

i. Task Description
According to Stevens & Levi (2013), task description describes the performance and expected behaviour from the assignment given such as in the form of paper, posters or performances, or even through observations such as laboratory regulations and their involvement in any activities.

ii. Scale
An effective rubric has three to five criteria, as mentioned by Popham (1997). Meanwhile, Stevens and Levi (2013, ms 11) agreed that a minimum of three scales should be selected for a rubric. Likewise, Wolf and Stevens (2007) pointed out that a good rubric should have less than six criteria.

iii. Dimension
One component that can be measured is dimension (Stevens & Levi, 2013) and this can be seen through assessments of writing skills; form of dimensions are accuracy, grammar, syntax, and spelling (Lovorn, Michael & Rezaei, 2011). Meanwhile, for conversation skills, measured dimensions are attention, exposure, expression and coordination (Sultana et al., 2012).

iv. Description of Dimension, Assessment Criteria or Descriptor
Descriptors refer to the explanations of dimensions or assessment criteria to identify the dimension descriptions (Montgomery, 2000). Stevens and Levi (2013) stated that it is advisable for a description of each measured dimension to be written in simple and ordinary texts. The quality of the descriptor selected and the subject matter experts involved during the development phase are the deciding factors of an efficient rubric (Garfolo et al., 2016). Rubric will become more flexible and personal if the students can see an addition to their scores for every dimension. This can help and motivate them to achieve the level that they yearn for. However, this is also a weakness due to the fact that the lecturers might have to take some time to explain the remaining scores that they miss (Stevens & Levi, 2013).

A rubric does not have a specific format. Therefore, a basic format is briefly explained in Table 1, Table 2 and Table 3 (Stevens & Levi, 2013; Wolf, Connelly, & Komara, 2008) below.
Rubrics Development Models

The first step in developing a rubric is to set up the quality or performance standard (Moskal, 2003) followed by determining the lowest performance level, highest performance level, median performance level and also the number of categories. Andrade (1997) further developed a rubric initiated by teachers where he began by looking at several work models of students from previous classes to determine the strengths and weaknesses of a task. Then, he listed down some criteria, set the quality grade and asked the students to use the rubric as a tool for self-assessment or peer evaluation, rechecked the rubric and then the teachers use the rubric. Two detailed development models are explained in the next section.
Stevens and Levi’s (2013) Model

As elaborated by Stevens dan Levi (2013, ms 29-30), the roles of developer (lecturers) and users (students) are determined through five development models and four level of development phases. The development models are presentation model, feedback model, Post-It model, 4x4 Model and Pass-The-Hat model.

Next, the four steps to develop a rubric are as follows:

i. Step 1: Reflection
Reflection is not conducted only for a specific assignment but it should be conducted for the objective of the whole course. Rubric can be developed by asking questions like “what are the skills needed to successfully complete an assignment?”, “Does the assignment provide the expected evidence from the students?”, “What is the anticipated highest performance from an assignment?”, and “What is the expected performance that students cannot achieve?”.

ii. Step 2: Listing
Details of an assignment need to be listed to help in achieving the intended learning objectives. The factors that come into play when listing out the details are learning outcomes according to the level of study, skills, task format and learning goals. As suggested by Montgomery (2000), in order to reduce the difference in interpretation among rubric users, a clear and specific language should be used in the description process.

iii. Step 3: Grouping and Labeling
This step requires developers to collect similar performance expectations in the same group, and then label them correctly. However, in this category, learning outcomes are not clearly stated as compared to individual descriptors. This category is the dimension part of the rubric.

iv. Step 4: Applying
An action of moving the developed group or list into a grid in levels. In this step, the performance expectation group is labeled as a “dimension” and it is located on the left column of the rubric grid.

For the presentation model, it is prepared by lecturers by following all steps of (i) to (iii). During step (iv), students and lecturers use the rubric to ask questions and have some reflections on their understandings. It is different for feedback model, where steps (i) to (iii) are prepared by lecturers, but at step (iv), lecturers and students provide feedback and allow changes to the rubric for clarity. For Pass-The-Hat model, rubric is prepared by lecturers during step (i) and then the criteria in step (ii) to (iv) are listed together with the students. Meanwhile, for Post-It model, lecturers prepare step (i) while students list down the criteria at step (ii), followed by the involvement of both lecturers and students in steps (iii) and (iv) in completing the final rubric.

For 4x4 rubric model, students take over almost all of the steps. The rubric development model is selected based on the suitability of the size of the class, and the rubric covers all levels.
including graduate and postgraduate students. The roles of lecturers and students in the development of rubrics at each step are summarized in Table 4.

Table 4
The Role of Lecturers and Students in the Development of Rubrics at Each Step

<table>
<thead>
<tr>
<th>Rubric Development Model</th>
<th>Step 1 Reflection</th>
<th>Step 2 Listing</th>
<th>Step 3 Grouping and Labeling</th>
<th>Step 4 Applying</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation</td>
<td>Lectures</td>
<td>Lecturers</td>
<td>Lecturers</td>
<td>Lecturers and students ask questions and reflect on their understanding.</td>
</tr>
<tr>
<td>Feedback</td>
<td>Lecturer</td>
<td>Lecturer</td>
<td>Lecturer</td>
<td>Lecturers and students edit for clarity.</td>
</tr>
<tr>
<td>Pass-the-hat</td>
<td>Lecturer /Student</td>
<td>Lecturer and students gather the contributed criteria list</td>
<td>Lecturer and students develop the final rubric</td>
<td></td>
</tr>
<tr>
<td>Post-it</td>
<td>Lecturer</td>
<td>Student</td>
<td>Lecturer and students facilitate the collection of the contributed criteria list</td>
<td>Lecturer and students develop the final rubric</td>
</tr>
<tr>
<td>4 x 4</td>
<td>Lecturer /Student</td>
<td>Student</td>
<td>Students develop the final rubric</td>
<td></td>
</tr>
</tbody>
</table>

**D Churches’s Model (2015)**
In this rubric development 4D Churches’s model, four steps are involved which are define, design, do and debrief.

a. Assignment Development (Define)
At this stage, the assignment is not yet determined so the instructor will identify the rubric’s main objectives, elements or components, clarify the tasks required, design the assignment and identify how rubrics can match the task or learning outcomes.

b. i. Assessment Mode (Design)
Assessment is determined to be formative or summative at this stage. The rubric will be used by the rubric developer to identify how frequent rubrics are used, what aspects of feedback are given and are rubrics used for lecturer assessments, peer assessment or self-assessment.

b.ii. Assessment Design (Design)
The rubric style and layouts to be used will be shown in the design phase in order to fulfill the rubric components, such as task description, scale, dimension and dimension descriptions.

c. Rubric Development (Do)
The rubric is developed with the quality of performance expectations along with the appropriate styles and layouts after the completion of the definition and design phase. The involvement of students in rubric development is supported by Churches due to the fact that it is a good learning process.
d. Use and Evaluation (*Debrief*)
The suitability of rubric to users and whether it meets the objectives are assessed after rubric has been used. The use of rubric is consisted of:

- Does the rubric accurately assess the components or elements?
- Does the rubric provide appropriate feedback?
- Does the rubric provide opportunities for improvement?
- Is the rubric easy to use?
- Is the rubric easy to understand?
- Is the rubric clear, compact and suitable for a range of age?
- Is the rubric in line with the goal?
- Overall, is the rubric suitable to use?

(Churches, 2015)

**Assessment and Rubric Scoring**
An achievement cannot be well defined through the numbering of measurements. Thus, this problem can be overcome by using rubric scoring due to the existence of dimension descriptions. The description in the rubric satisfies the rating scale characteristic of an instrument (Company et al., 2017), which is known as the basis of scoring.

Karkehabadi (2013) stated that if rubric is used for assessment for learning purpose, the evaluation can be conducted using qualitative and quantitative approaches. On the other hand, if rubric is used for assessment of learning or summative purposes, the evaluation can be done using a quantitative approach.

The checklist is the same as an analytic scoring rubric to allow a separate assessment to be carried out according to dimensions as described by Moskal (2000). However, it is preferable to use a holistic scoring rubric if there is an overlap between the criteria set to evaluate the various dimensions, where criteria are considered as a combination of a single descriptive scale.

Various appropriate grading methods for formative assessment through rubric have been suggested by Stevens and Levi (2013, p. 75), as indicated in Table 5. Formative grading method is not formed using quantity aspect as it is not a priority. However, summative grading method is still done by scoring at each level of the overall rating, although the rubric is often used as a formative assessment and self-assessment tools. Rubric users can apply the percentages or cumulative grades using summative grading method (Company et al., 2017).

Janssen, Meier dan Trace (2015) applied a composite profile that contains a range of scores, criteria description, and comments at the end of each construct (Figure 1).
Analysis of threshold value determination for achievement leveling can be done using summative grading (Janssen et al., 2015). For example, parameter estimation, $\delta$ is used to estimate the parameters of student’s ability as well as to determine the level as discussed in Sadler’s (2009) study (Figure 2). Determination of Rasch-Andrich Threshold is calculated using Rasch modeling (Linacre, 2006).

For cut-off score determination, the levels are divided into achievement in grading, and the best number of categories which is less than three or four categories (Zieky & Perie, 2006, ms 4). Then, by using the proficiency level as applied in the No Child Left Behind (NCLB) test, it is categorized into four levels, namely ‘below basic’, ‘basic’, ‘proficient’ and ‘advanced’ (Bejar, 2008, ms 1). According to Zieky and Perie (2006), cut-off score categorization depends on subject-matter-content and policy changes. Rasch modeling which is known as “Rasch-Andrich threshold”, “step calibration” or “tau”, $\tau$ is the point where a latent variable has a probability of occurrence in category $j$, equivalent to the probability for a rating in category $j-1$ (Linacre, 2012).

**Conclusion**

Even though a guideline for rubric construction has been given, the specific shape of rubric can be modified according to the assessment forms. Hence, the use of the rubric should be adjusted accordingly with the objective of the assessment.
The developers and users are the ones who gain the most benefits from the rubric. Thus, it is in line with its development goal. In fact, the use of rubrics by students and teachers will be easier through various technology platforms. Therefore, the rubric should be accepted psychometrically as a measuring instrument. The next stage in rubric construction is to determine its psychometrics measurement of validity and reliability, will be discussed in the next paper.
<table>
<thead>
<tr>
<th>Rubric Layout</th>
<th>Marking Method</th>
<th>Grading Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Three to five rating levels with checkmarks</td>
<td>Checkmark</td>
<td></td>
</tr>
<tr>
<td></td>
<td><img src="image1.png" alt="Rubric Layout" /></td>
<td><img src="image2.png" alt="Grading Method" /></td>
</tr>
<tr>
<td>ii. Three to five rating levels</td>
<td>Circled text</td>
<td></td>
</tr>
<tr>
<td></td>
<td><img src="image3.png" alt="Rubric Layout" /></td>
<td><img src="image4.png" alt="Grading Method" /></td>
</tr>
<tr>
<td>ii. Rubric scoring</td>
<td>Narrative feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td><img src="image5.png" alt="Rubric Layout" /></td>
<td><img src="image6.png" alt="Grading Method" /></td>
</tr>
<tr>
<td>Rubric Layout</td>
<td>Marking Method</td>
<td>Grading Method</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Task Description: Prepare a one-hour presentation on an assigned film. You are expected to discuss how the film relates to the political, economic, or cultural aspects of the historical period it claims to depict and also the cinematography of that era. You may also discuss it in terms of film theory if you wish.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HIST Black Rain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimensions</td>
<td>Criteria</td>
<td>Comments</td>
</tr>
<tr>
<td>Introduction</td>
<td>The introduction tells the audience exactly what to expect in terms of how the speaker feels about the movie, what theories and theoretical framework(s) he or she will introduce, and what conclusions he or she will draw.</td>
<td>Allotted time should be included PowerPoint is to be the only projection.</td>
</tr>
<tr>
<td>Organization</td>
<td>The pretransitn is required to create a logical argument, and all that topics that need to be discussed together are presented together.</td>
<td>Well-ordered flow; thematic and cohesive organization; connections.</td>
</tr>
<tr>
<td>Content</td>
<td>The presenter discusses the main historical issues raised by the film and how other film scholars and historians have dealt with these issues both within the film and in general. The presentation explicates where he or she stands on these issues, which theories he or she finds most useful, and why.</td>
<td>Discuss: - Social, political, cultural and religious - Industry and community patterns - Setting and village life - Japan's role and position.</td>
</tr>
<tr>
<td>Evidence</td>
<td>The presenter includes sufficient, detailed examples from the film and other sources to support his or her analyses.</td>
<td>Appropriately uses filmic and supplementary sources.</td>
</tr>
<tr>
<td>Analysis</td>
<td>The presenter uses or gives evidence to support a concise, coherent analysis of how the film fits in or does not contribute to our understanding of World War II.</td>
<td>Thorough. A critical, impartial analysis outlining all elements listed previously on a scholarly basis despite the speaker's emotional or biased stance.</td>
</tr>
<tr>
<td>Presentation</td>
<td>The presenter speaks clearly, loudly enough to be heard, but not too loudly; used appropriate, effective gestures and body language; and maintained eye contact with the class. Also used visual aid, if used, was technically sound (no poweroutage, loudspeakers, inappropriate, and offensive to the presentation.</td>
<td></td>
</tr>
</tbody>
</table>

1. Post-traumatic stress disorder
2. Black and white as opposed to color film
Acknowledgement
The authors acknowledged the Niche Research Grant Scheme (NRGS) Project (Assessment), Universiti Pendidikan Sultan Idris (2014-00001-107-82-3) and the Ministry of Education Malaysia for the completion of this article.

Corresponding Author
Name: Azliza Muhammad
Malaysia
Email: muhammadazliza@gmail.com

References


---