Music Educators Journal

Understanding and Developing Rubrics for Music Performance Assessment
Brian C. Wesolowski

The online version of this article can be found at:
http://mej.sagepub.com/content/98/3/36

Published by:
$SAGE

http://www.sagepublications.com

On behalf of:

National Association for Music Education

National Association for Music Education

Additional services and information for *Music Educators Journal* can be found at:

Email Alerts: http://mej.sagepub.com/cgi/alerts

Subscriptions: http://mej.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Mar 7, 2012

What is This?

by Brian C. Wesolowski

Understanding and Developing Rubrics for Music Performance Assessment

Abstract: A primary difficulty with music performance assessment is managing its subjective nature. To help improve objectivity, rubrics can be used to develop a set of guidelines for clearly assessing student performance. Moreover, rubrics serve as documentation for student achievement that provides music teachers with a written form of accountability. This article examines the complexities of music performance assessment and provides an argument for the benefit of rubrics in the assessment process. In addition, discussion includes an overview of the various types of rubrics as well as suggestions for choosing and writing rubrics to assess musical performances.

Keywords: accountability, assessment, evaluation, rubric, performance, professional development, teacher education

Assessment rubrics can give music teachers and their students a clear and fair measure of progress and mastery in the classroom, studio, or rehearsal.

In the United States today, assessment has become one of the most important and pervasive topics in music education.[1] Assessment can be defined as the collection, analysis, interpretation, and application of information about student performance or program effectiveness in order to make educational decisions.[2] It consists of two important components: (1) measurement of student performance and (2) evaluation of data.[3] Educational reform and related matters of accountability have provided grounds for the widespread demand for new assessment strategies.[4]

Now more than ever, teachers of all grade levels find themselves in situations in which a thorough documentation of student performance is necessary. According to Edward Asmus, the fundamental rationale for assessment is that "the more a teacher and student know about the student's learning, the more effective the teacher can be in facilitating learning and the more effective the student can be in acquiring the learning."[5] In addition, the documentation of student performance enhances parents' abilities to support and assist in the process of student learning and achievement.[6]

Numerous textbooks have been written to help classroom teachers develop their ability to implement strategies for performance assessment. These books, however, tend to approach assessment in terms of objective

Brian C. Wesolowski is a doctoral candidate and graduate teaching assistant at the University of Miami, Coral Gables, Florida. His research interests include music performance assessment, scale development for performance measurement and evaluation, and the microstructure, perception, and pedagogy of jazz rhythm. He can be contacted at BCWesolowski@gmail.com.

testing, in which there is only one clear, correct answer. Music, on the other hand, is a discipline that embraces expressive decisions and divergence of response. By nature, music educators deal with performance assessment on a daily basis, using techniques to recognize, diagnose, and communicate methods for improvement in individual or ensemble performance. This yields numerous opportunities to assess student performance achievement. However, many music educators have not developed a mechanism to document student learning and developmental progress, or they have not tied their individual student assessment to their grading system. Furthermore, many music teachers rely heavily on nonmusic criteria, such as behavior, attitude, attendance, and participation to determine their grades.[7] Often, students are not receiving a clear indication of their individual achievement, how they can improve, or the teacher's expectation of their individual performance. Music educators need to reevaluate their assessment strategies. The inclusion of more formative assessment methods directly tied to student achievement must be explored. For music educators to be successful at music performance assessment, they must be not only prescriptive in evaluating their individual students and ensembles but also willing to assess and improve their methods of teaching and communicating.

Complexities of Assessment

Music performance assessment is a complex process. A primary difficulty with performance assessment is managing its subjective nature.[8] The testing and assessment of music will always be affected by aesthetic value judgments. Human reasoning is used to establish test criteria, determine the method of testing, determine the method of scoring, and interpret scores. The combination of this subjectivity with the ephemeral nature of music makes music performance assessment an especially arduous task. David J. Boyle and Rudolf E. Radocy suggest that assessments can be significantly improved through assessment measures with improved measures of objectivity.[9]

A more valid form of assessment may be used by building objective measures based on aesthetic value judgments. Rubrics can be used to develop a set of guidelines that can help educators clearly assess students' work. Rubrics not only aid in increasing the objectivity of teacher grading, but they also can be tailored to assess ensemble members on an individual basis.

Another complexity with music performance assessment is the use of summative and diagnostic assessments as a means to document individual student achievement.[10] Summative and diagnostic assessments may not always provide the necessary feedback to improve instruction. These assessments relate learning to the quantification of errors in any given performance. By using only summative and diagnostic assessment methods, a teacher cannot get a clear indication of whether the student is truly learning about music or only learning the teacher's parameters for the tolerance of mistakes.[11] If assessments in music performance are implemented with a purely summative or diagnostic purpose, we lose the richness of instructional value that could be possible. By implementing more formative methods of assessment, such as the rubric, music educators can better monitor and improve student learning as well as shape their instruction in a tangible, sequential manner in response to what they discover.

Rating Scales

Research in music education has yielded two primary rating scales for music performance assessment: (1) Likert-type scales and (2) criteria-specific scales.[12] A Likert-type scale is a psychometric scale in which the evaluator is asked to respond to a list of item statements by his or her level of agreement or disagreement. Likert-type scales provide clear evidence of an evaluator's judgment of a specific performance.[13] However, they lack the individualized detail of formative assessments and do not offer information about the mastery of performance aspects.[14] The use of statistically validated Likert-type scales can be found in research studies of

solo performance assessment as well as ensemble performance assessment.[15]

Criteria-specific performance scales are based on written, objective statements that describe various performance attributes. These objective statements offer more information to the student than assessments using Likert-type scale responses because they offer insight into proficiency levels. The scales are generally graded in a dichotomous manner. Researchers have traditionally used these scales in either additive approaches or continuous approaches.[16] Scales involving an additive approach demonstrate a list of criteria that becomes more developed as the list continues. Scales using a continuous approach demonstrate a list of criteria that assumes the form of an equally weighted checklist. The benefits of criteria-specific performance scales are that they are able to assess very specific levels of performance aptitude accurately and reliably. However, adjudicators may find difficulty in judging music performances based on single, generalized, objective statements.[17] Also, these scales do not offer any type of quality judgment or convey the level of achievement. There is only a judgment of "present" or "absent" according to the specified criteria in the checklist.

The rubric is a form of a criteria-specific performance scale. It is a set of scoring criteria used to determine the achievement level of a student's performance on assigned tasks. A rubric divides a task into constituent parts and offers detailed descriptions of the performance levels for each part.[18] The descriptions are written so students are able to learn what must be done to improve their performances in the future. Because it helps teachers directly assess performance experiences, a rubric is a tool for providing authentic assessment. Educational literature tends to make the rubric selection process overly dramatic and complicated. Each branch of education has its own methodology for adaptations of types of rubrics and their usage. Furthermore, the overlap between these branches can lead to more complexity and ambiguity. The unclear nature of rubric writing has left many educators with a vague understanding of their operational

use and advantages. Recently, there has been a surge of published research dealing with the statistical validation of rubrics.[19] In spite of this, more guidance into rubric construction is needed.

Benefits of Using Rubrics

Rubrics offer many advantages in the music performance medium. Rubrics can provide the following:

1. Clear levels of accomplishment by defining tangible measures of individual achievement

2. Clear indications of what students need to accomplish in the future to improve their individual performance

3. A learner-centered approach to performing, learning, and assessing

4. A bridge between student learning and teacher expectation

5. Versatility in adapting to meet the needs of a specific curriculum, student age, ability level, style of music, and type of ensemble

6. A valid and reliable form of individualized assessment and documentation of teacher accountability

7. A quantitative means for evaluating and scoring qualitative, performance-based tasks

8. A means for clearly implementing content standards and course objectives into the assessment process

9. Valuable information for parents on their child's progress and needs for improvement

Types of Rubrics

There are two main categories of rubrics: holistic and analytic.[20] Holistic rubrics provide a single score based on an overall assessment of a music performance. The evaluator matches the descriptors of the scale to his or her overall impression of the performance.

Generally, a holistic rubric is written in a manner that is generic and simple enough to adapt to other performance situations. A holistic rubric written to assess a concert band should also be able to serve as an assessment tool for an orchestra or a jazz band. A holistic rubric written to assess a solo clarinet performance should also be general enough to use for other solo instrumental performances. An advantage of holistic rubrics is that they are easy and fast to use. However, they do not provide detailed information on an overall performance assessment. The score of the holistic rubric will not provide the student with specific feedback for the teacher's choice of grading. Figure 1 is an example of a task-specific, holistic rubric designed as a quarterly assessment for individual student sight-reading within the ensemble.

The opposite of a holistic rubric is the analytic rubric. An analytic rubric contains more than one dimension of evaluative criteria. The multiple criteria are matched with multiple descriptors and the teacher's feedback, and scoring is based on each of these individual dimensions. Because of the assessment by multiple criteria, the analytic rubric provides more information than does the holistic rubric. The analytic rubric consists of multiple scales, thereby providing multiple sets of scores. A benefit of analytic rubrics is the wealth of specific, individualized assessment information that can be of great value to students, parents, and teachers. These rubrics move beyond basic, generic descriptions of tasks. In this type of rubric, a student's performance can vary across performance levels because complex concepts are broken into constituent parts. Students may master one area but perform in an average or below-average manner in another. As an example, a student might have achieved an accomplished level of performing with good intonation, but may still be developing accuracy in rhythm. This variability across dimensions allows students, parents, and teachers to isolate the strengths and weaknesses of the overall performance assessment. Analytic rubrics are generally task specific in nature; however, they can be generic or a mixture of generic and task specific as well. Figure 2 shows an example of a task-specific, analytic rubric designed as a weekly,

individual student assessment for ensemble music preparation.

Creating Your Own Rubric

Define the Focus, Purpose, and Objectives of the Assessment

Before creating a rubric, the teacher needs to know each learning goal and be able to define levels of accomplishment. This initial reflection process should not only entail the specific reasoning for the particular assessment but also include attention to the overall performance structure, the needs of the specific students being assessed, the expectations of what is to be accomplished, and the students' prior knowledge and skill.

The focus and purpose of the rubric represented in Figure 2 is to assess each member of an advanced high school wind ensemble. Specifically, it is intended to provide frequent feedback about each student's musical development and foster a better sense of individual accountability for each student's musical choices within the context of ensemble performance.

Define the Performance Criteria and Learning Outcomes

The performance criteria should be clearly defined, with no ambiguity across the various criteria domains. Each criterion listed should be an important learning outcome for a high-quality performance and understood by the student. Choose criteria that reflect your teaching goals. The rubric represented in Figure 2 focuses on four learning outcomes:

1. rhythmic fluidity within the melodic line,

2. tone control within varying registers,

3. control of intonation and self-demonstrated intonation adjustment, and

4. consistency of focus within the rehearsal setting.

Each of the four performance criteria focuses on the specific musical task

FIGURE 1

Example of a Task-Specific, Holistic Rubric Designed as a Quarterly Assessment of Individual Student Sight-Reading within the Ensemble

Quarterly Individual Sight-Reading Performance Assessment
4 points (Accomplished) *Grade: A*
Intonation and balance: Student plays with good overall intonation and demonstrates ability to match pitch quickly and fluently.
Fluency of reading notation in ensemble context: Reading is fluid with very few to no pauses.
Phrasing: Interpretation of melodic contour, note emphasis, and note duration are performed consistently in a proper, idiomatic style.
3 points (Proficient) *Grade: B*
Intonation: Student plays with moderately-good overall intonation and demonstrates some difficulty matching pitch.
Fluency of reading notation in ensemble context: Reading is moderately fluid with occasional pauses.
Phrasing: Interpretation of melodic contour, note emphasis, and note duration are performed somewhat consistently in a proper, idiomatic style.
2 points (Developing) *Grade: C*
Intonation: Student does not consistently play with good overall intonation and demonstrates difficulty in matching pitch.
Fluency of reading notation in ensemble context: Reading lacks fluidity with frequent pauses with little coherency.
Phrasing: Interpretation of melodic contour, note emphasis, and note duration are performed mostly inconsistent of the idiomatic style.
1 point (Beginning) *Grade: D*
Intonation: Student does not play with good intonation and demonstrates little to no awareness in matching intonation.
Fluency of reading notation in ensemble context: Reading lacks fluidity and is disjunct with no coherency.
Phrasing: Interpretation of melodic contour, note emphasis, and note duration are performed inconsistently in the idiomatic style.
Comments: *You have come a long way! Continue to work on your intonation exercises with a tuner, particularly in the upper register! It continues to be a bit sharp at times.*　　Score: *3*

that the teacher finds to be consistently troublesome, as well as the student's responses to teacher prompting.

Determine the Type of Rubric for Your Assessment

Is the assessment formative or summative? Are you teaching a new skill or further developing a previously-taught skill? Look at the complexity of the skill and how many component parts are to be assessed. In addition, consider how closely related the learning outcomes are to each other. If there is a certain level of overlap, a holistic rubric may be a

better fit. How much time do you have to invest in the assessment process? Analytic rubrics take more time to develop and grade. The rubric demonstrated in Figure 2 is for formative assessment purposes. It is intended as a weekly assessment for each ensemble member's music preparation. The rubric has two purposes: (1) to continue to develop basic musical skills, such as tone, intonation, and rhythmic fluidity, and (2) to foster a new skill of individual accountability for their musical choices and individual focus within the rehearsal. An analytic rubric was chosen because there is diversity in the learning outcomes, and the teacher concluded that

the time invested to fill out the rubrics was worth the reward.

Define the Range and Degrees of Proficiency of Performance Scale Levels

The degree of proficiency describes how well the dimension has been performed.

There are unlimited labels that can serve to categorize the levels of proficiency achieved by the students, for example, (1) beginning, (2) developing, (3) accomplished, and (4) exemplary. The majority of performance assessment rubrics tend to contain three to five categories.

FIGURE 2

Example of a Task-Specific, Analytic Rubric Created as a Weekly, Individual Music-Preparation Assessment for an Advanced High School Wind Ensemble

Learning Outcome	Beginning	Developing	Accomplished	Exemplary	Score
		Weekly Individual Music Preparation Assessment			
Rhythmic fluidity within the melodic line	☐ 1 pt. Student plays with little or no fluidity.	☐ 2 pts. Student plays with some fluidity, but does not improve with coaching.	☑ 3 pts. Student plays with some fluidity and improves with coaching.	☐ 4 pts. Student plays with fluidity and self-adjusts.	3
Tone control within varying registers	☐ 1 pt. Student plays with little or no control.	☐ 2 pts. Student plays with some control, but has not shown improvement.	☐ 3 pts. Student plays with some control and shows improvement with coaching.	☑ 4 pts. Student plays with control and self-adjusts.	4
Control of intonation and demonstrated intonation adjustment	☐ 1 pt. Student needs considerable attention to intonation problems.	☐ 2 pts. Student has some problems with intonation, but does not improve with coaching.	☐ 3 pts. Student has some problems with intonation and shows improvement with coaching.	☑ 4 pts. Student plays in tune and self-adjusts.	4
Consistency of focus in rehearsal setting	☐ 1 pt. Overall performance is almost always inconsistent.	☐ 2 pts. Overall performance is generally inconsistent.	3 pts. Overall performance is generally consistent.	☐ 4 pts. Overall performance is consistent.	3

Comments: *Continue to work on your fluidity in performance: the connecting and musical shaping of phrases, evenness of notes (in rapid passages), and creating a relaxed feeling during technical passages so as not to rush. Overall, you are gaining much independence in making musical decisions and adjustments. Continue working hard! Bravo!*

Sum score:

14—A

Explanation of Scoring: Each proficiency level is scored according to the label: Exemplary (4 points), Accomplished (3 points), Developing (2 points), and Beginning (1 point). Students are graded according to each criterion under the level and the total should be added. The maximum score for this rubric is 16 points. The scoring cut points are as follows:

A: 14–16
B: 11–13
C: 8–10
D: 4–7

The rubric demonstrated in Figure 2 includes four degrees of proficiency: (1) emerging, (2) progressing, (3) partial mastery, and (4) mastery. The specific terminology was chosen to accommodate the level of the students in the ensemble. For example, choosing the term *beginning* for the first level would not fit the skill level of an advanced high school wind ensemble. In addition, the term *mastery* may provide a goal to achieve for an older student who can fully comprehend the term.

Define Appropriate Task Expectations and Meaningful Descriptors for Each Criterion Performance Level

The totality of the descriptors provides a comprehensive summary of what is being assessed. The descriptors should be written as clearly and concisely as possible. Avoid any vernacular or terminology that is superfluous in nature. Write descriptors for continuity between levels of performance in each category. The descriptors should define a continuum of quality throughout each category. Be sure that each descriptor has a clear sense of flow between levels. The descriptors should be detailed enough to limit subjectivity yet concise enough to avoid confusion or ambiguity. Edwin Gordon states that the more descriptors that are included for each dimension, the more reliable the rubric will become; however, the number should not exceed five.[21]

The rubric demonstrated in Figure 2 contains one descriptor for each level. The descriptors were specifically written as matched pairs: "Emerging/Mastery" and "Progressing/Partial Mastery." The most important descriptions are contained in the two middle descriptors. The *progressing* category demonstrates that the student is getting better at the skill but does not respond independently to the prompting of the teacher. The *partial mastery* category demonstrates that the student is getting better at the skill and demonstrates improvement with prompting. These two descriptions, skill demonstration and response to prompting, are found throughout all four learning outcomes.

Choose an Appropriate Scoring Scale with Clearly Defined Cut Points

The levels of a rubric provide a reliable structure to which a teacher can assign numeric values for the performance area. In terms of teacher accountability, it is important to construct a logical and easy-to-understand scoring guide that correlates to the entirety of the performance assessment. The method and style of scoring may vary from one assessment to the other. There are generally three methods to choose from for grade construction: (1) assign each scale level a point value, and sum each descriptor to reach a defined grade according to a predetermined scoring guide; (2) assign each scale level a point value while weighting each descriptor according to importance within the total assessment; or (3) assign a letter or numeric value for each level. In each case, a clearly defined scoring method must be labeled and maintained throughout the assessment process.

The rubric demonstrated in Figure 2 demonstrates a four-letter grading scale. Each degree of proficiency is assigned a point value. The learning outcomes are graded accordingly and added to form a sum score. The creation of the grading scale was constructed in three stages: (1) the lowest possible outcome (in this example, 4) was specifically chosen to be represented by the letter grade of D; (2) letter grades were chosen (in this example, four grades: A, B, C, and D); and (3) each letter grade was divided into equal divisions.

A Valuable Tool

This article is a brief introduction to the understanding and developing of rubrics for music performance assessment. Many variations of rubrics exist. As a teacher becomes increasingly comfortable with rubric development, more variations may be developed to accommodate specific ensembles, students, student needs, and teaching styles. The outcomes that result from the implementation of assessment tools such as the rubric can enlighten the teaching and learning processes. Rubrics can serve as a valuable assessment tool for music educators to assist them in determining the overall effectiveness of the educational process.

NOTES

1. Edward Asmus, "Considerations for Teaching Music Education Assessment," in *Proceedings of the Tenth International Symposium of the Research Alliance of Institutes for Music Education,* ed. Sven-Erik Holgersen and Frede V. Nielsen (Copenhagen: RAIME, 2010), 27.

2. Edward Asmus, "Music Assessment Concepts," *Music Educators Journal* 86, no. 2 (1999): 21.

3. David J. Boyle and Rudolf E. Radocy, *Measurement and Evaluation of Musical Experiences* (New York: Schirmer Books, 1987).

4. Asmus, "Music Assessment Concepts," 19.

5. Asmus, "Considerations," 27.

6. Anthony W. Jackson and Gayle A. Davis, *Turning Points 2000: Educating Adolescents in the 21st Century* (New York: Teachers College Press, 2000).

7. Claire W. McCoy, "Grading Students in Performing Groups: A Comparison of Principals' Recommendations with Directors' Practices," *Journal of Research in Music Education* 39 (1991): 181–90.

8. Rudolf E. Radocy, "On Quantifying the Uncountable in Musical Behavior," *Bulletin of the Council for Research in Music Education* 88 (1986): 22–31.

9. Boyle and Radocy, *Measurement.*

10. Thomas W. Goolsby, "Assessment in Instrumental Music Education," *Music Educators Journal* 86, no. 2 (1999): 31–32.

11. Ibid., 32.

12. T. Clark Saunders and John M. Holahan, "Criteria-Specific Rating Scales in the Evaluation of High School Instrumental Performance," *Journal of Research in Music Education* 45, no. 2 (1997): 260–61.

13. Boyle and Radocy, *Measurement.*

14. Saunders and Holahan, "Criteria-Specific Rating Scales," 259–72.

15. Harold F. Abeles, "Development and Validation of a Clarinet Performance Adjudication Scale," *Journal of Research in Music Education* 21, no. 3 (1973): 246–55; Martin J. Bergee, "An Objectively Constructed Rating Scale for Euphonium and Tuba Performance,"

Dialogue in Instrumental Music Education 13, no. 2 (1989): 65–81; Robert H. Horowitz, "The Development of a Rating Scale for Jazz Guitar Improvisation Performance" (PhD diss., Columbia University Teachers College, 1994); Jon P. Nichols, "A Factor Analysis Approach to the Development of a Rating Scale for Snare Drum Performance" (PhD diss., University of Iowa, 1985); Brian E. Russell, "The Development of a Guitar Performance Rating Scale Using a Facet-Factorial Approach," *Bulletin of the Council for Research in Music Education* 184 (2010): 21–34; Stephen F. Zdzinski and Gail V. Barnes, "Development and Validation of a String Performance Rating Scale," *Journal of Research in Music Education* 50, no. 3 (2002): 245–55; John M. Cooksey, "A Facet-Factorial Approach to Rating High School Choral Music Performance," *Journal of Research in Music Education* 25, no. 2 (1977): 100–114; Charles B. DCamp, "An Application of the Facet-Factorial Approach to Scale Construction in the Development of a Rating Scale for High School Band Performance" (PhD diss., University of Iowa, 1980); and Bret P. Smith and Gail V. Barnes, "Development and Validation of an Orchestra Performance Rating Scale," *Journal of Research in Music Education* 55, no. 3 (2007): 268–80.

16. Kenneth U. Gutsch, "Evaluation in Instrumental Music Performance: An Individual Approach," *Bulletin of the Council for Research in Music Education* 4 (1965): 21–29; Robert L. Kidd, "The Construction and Validation of a Scale of Trombone Performance Skills" (PhD diss., University of Illinois at Urbana–Champaign, 1975); and Saunders and Holahan, "Criteria-Specific Rating Scales."

17. Abeles, "Development and Validation."

18. Danielle D. Stevens and Antonia A. Levi, *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning* (Sterling, VA: Stylus, 2005).

19. Charles R. Ciorba and Neal Y. Smith, "Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric," *Journal of Research in Music Education* 57, no. 1 (2009): 5–15; Marvin E. Latimer Jr., Martin J. Bergee, and Mary L. Cohen, "Reliability and Perceived Pedagogical Utility of a Weighted Music Performance Assessment Rubric," *Journal of Research in Music Education* 58, no. 2 (2010): 168–83; and Charles E. Norris and James D. Borst, "An Examination of the Reliabilities of Two Choral Festival Adjudication Forms," *Journal of Research in Music Education* 55, no. 3 (2007): 237–51.

20. Audrey M. Quinlan, "*A Complete Guide to Rubrics: Assessment Made Easy for Teachers, K–College* (Lanham, MD: Rowman & Littlefield Education, 2006).

21. Edwin Gordon, *Rating Scales and Their Uses for Evaluating Achievement in Music Performance* (Chicago: GIA, 2002).