Group Project Report:

Rubric for High School Choir Performance

Anthony DiTaranto

A&HM 5002: Assessment Strategies for the Arts

Dr. Harold Abeles

Teachers College, Columbia University

Introduction

Assessing musical performance can be one of the more difficult and sensitive areas of music education. By collaborating on this rubric, our group intends to bring objectivity with subjectivity and specificity with sensitivity to assessing performance. The "figs" group selected high school choir to be the focus of our assessment. We designed a rubric based on the components of musical performance we agreed were the most valid: vocal technique, intonation, rhythm, style, diction, and overall musical performance. We defined four rubric progressions for each component and articulated criteria as objectively as possible. Three judges scored five performances (in the form of YouTube videos). We found our assessment to be reliable, but authenticity is called into question.

Factors Considered in Designing the Test

When selecting a type of group to evaluate, the group agreed on high school choir as it was the type of ensemble we had the greatest collective experience with. Using our own professional knowledge and example choral performance scoresheets as reference we each contributed ideas about what the salient features of choral performance are. The group agreed on categories relating to accuracy (intonation, vocal technique, diction, and rhythm), musicality (style), and expression (overall musical performance). Descriptions of each rubric progression were articulated with attention to objectivity, allowance for some subjectivity, and clear gradation between levels. Each category includes a description to help guide the listening of the judge to equivalent features. Each category has four levels, to eliminate the middle ground of five level scales. Each level also has a description (excelling, achieving, progressing, and beginning) chosen intentionally to communicate the level in a constructive manner. If a group receives low marks, we want them to be motivated to improve and know what details to work on; in short, we sought to create an "educative assessment" (Wiggins 2010).

Description of Student Sample

We selected the following five video performance samples to test our assessment:

Assessment A: https://youtu.be/EHAK263QoFo
Assessment B: https://youtu.be/tKslnZI_m9o
Assessment C: https://youtu.be/h9oWpKfFirs
Assessment D: https://youtu.be/qNbi3Ea61vM
Assessment E: https://youtu.be/SoqR54x-Lm0

Each is a high school choir of varying levels of selectivity and age. We believe that choosing groups with varying levels of experience will help illustrate the reliability of our assessment because we can expect a fairly consistent ranking from the three judges.

Summary of Analysis

Following is the data collected from our assessment. I evaluated our assessment using the Spearman Rho Rank-Difference Method and found our assessment to be reliable, with an average inter-judge reliability of 0.87.

JUDGES' SCORES

	Assessment A	Assessment B	Assessment C	Assessment D	Assessment E
Judge 1	18 (5)	23 (3)	24 (1)	21 (4)	24 (2)
Judge 2	16 (5)	21 (4)	22 (2)	22 (3)	23 (1)
Judge 3	14 (5)	18 (4)	24 (1)	21 (3)	24 (2)

RELIABILITY OF JUDGES 1 AND 2

	Assessment A	Assessment B	Assessment C	Assessment D	Assessment E
Judge 1	5	3	1	4	2
Judge 2	5	4	2	3	1
Difference	0	-1	-1	1	1
Differences^2	0	1	1	1	1
$\Sigma D^2 = 4$					
$r_s = 0.80$					

RELIABILITY OF JUDGES 2 AND 3

	Assessment A	Assessment B	Assessment C	Assessment D	Assessment E
Judge 2	5	4	2	3	1
Judge 3	5	4	1	3	2
Difference	0	0	1	0	-1
Differences^2	0	0	1	0	1
$\Sigma D^2 = 2$					
$r_s = 0.90$					

RELIABILITY OF JUDGES 1 AND 3

	Assessment A	Assessment B	Assessment C	Assessment D	Assessment E
Judge 1	5	3	1	4	2
Judge 3	5	4	1	3	2
Difference	0	-1	0	1	0
Differences^2	0	1	0	1	0
$\Sigma D^2 = 2$					
$r_s = 0.90$					

TOTAL INTER-JUDGE RELIABILITY

	Inter-judge Reliability
Judge 1-2	0.80
Judge 2-3	0.90
Judge 1-3	0.90
Average	**0.87**

Suggestions for Revisions

Although the assessment we designed is reliable, its authenticity needs further consideration. The rubric is intended to measure the quality of an artistic performance, which I believe is what arises when three types of goals are achieved: technical, musical, and emotional goals. Of the criteria evaluated in the rubric, four are technical (vocal technique, intonation,

rhythm, and diction), one is musical (style), and perhaps only one can be considered emotional (overall musical performance). To achieve a more complete depiction of the technical and affective aspects of musical performance, I would rethink the balance of the evaluated categories. I don't believe any technical elements should be sacrificed, but perhaps the rubric should be expanded with more musically and emotionally focused criteria. This would better represent the "fundamental thinking, production, and learning" in choral singing (Hafeli 2009).

The rubric could also be partially rewritten to include more "unique language" in the level descriptors (Abeles 2016a). Most follow the pattern of "most," "mostly," "somewhat," and "not." Richer language could provide better feedback and make the assessment more educative while encouraging judges to think more critically about each level.

Conclusions

Save for some possible expansion and rewriting, this rubric could be used reliably for high school level choral performances. It takes a wider perspective on performance than just technique and is very detailed in its descriptions. This allows it to provide actionable information to ensemble directors and students alike. The rubric was designed with all these features to create an assessment that is educative. It could be applied not only for post-performance, but pre-performance discussions and goal setting as well. Individual aspects could be explored in the context of class-level formative assessment to scaffold success in a summative performance assessment. Considering these factors along with its high reliability, this rubric is an excellent starting point for structuring choral achievement, and should be continually developed in the future.

	4 - Excelling	3 - Achieving	2 - Progressing	1 - Beginning
Vocal Technique Tone quality, resonance, and breath support	Tone is clear, resonant, and focused. Vocalists perform with breath support and control.	Tone is mostly clear and controlled. Few moments of unsupported breath or lack of projection.	Inconsistency in tone production. Moments of unfocused, unsupported tone.	Tone is inconsistent or unhealthy. Overall sound is unfocused and unsupported.
Intonation Accuracy to printed pitches	Intonation is consistently accurate throughout the vocal range and dynamic levels.	Intonation mostly accurate, but inconsistent at the extents of ranges and dynamic levels.	Intonation mostly accurate, but inconsistent through entire vocal ranges and dynamic levels.	Intonation rarely accurate through entire vocal range and all dynamic levels.
Rhythm Accuracy of note and rest values, duration, pulse, steadiness, correctness of meters	Rhythm is completely accurate, grounded in a steady pulse, pulse changes were coordinated together perfectly, and all entrances and cut offs were precise and correct.	Rhythm is mostly accurate, grounded in a steady pulse, pulse changes were coordinated together well, and most entrances and cut offs were precise and correct.	Rhythm is mostly accurate, steady pulse is somewhat unsure, pulse changes were somewhat coordinated together, and some entrances and cut offs were imprecise or incorrect.	Rhythm is mostly inaccurate, steady pulse is unsure, pulse changes were uncoordinated, and many entrances and cut offs were imprecise or incorrect.
Style Uniformity in dynamics and phrasing; interpretation and artistry	Dynamic changes are consistently coordinated together; conviction to and clear communication of the meaning of the content.	Dynamic changes are mostly coordinated together; moderate conviction to and clear communication of the meaning of the content.	Dynamic changes are unclear at times; little conviction to and clear communication of the meaning of the content.	Dynamic changes are not evident; no conviction to or clear communication of the meaning of the content.
Diction Clarity of consonants, consistency of vowels, pronunciation, clarity of text	Consonants are on time and together, vowels are consistent throughout the ensemble, and text is easily intelligible.	Consonants are mostly on time and together, vowels are mostly consistent throughout the ensemble, and text is mostly intelligible.	Consonants are sometimes on time and together, vowels are somewhat consistent throughout the ensemble, and text is somewhat intelligible.	Consonants are not on time and together, vowels are inconsistent throughout the ensemble, and text is unintelligible.
Overall Musical Performance Stage presence, artistry, appropriate appearance, poise, posture, general conduct, mannerisms, facial expression, memory, and audience engagement	Stage presence and concert etiquette are professional and appropriate. Ensemble watches the conductor, and presents a highly engaging performance.	Stage presence and concert etiquette are mostly professional and appropriate. Ensemble watches the conductor for the majority of the piece, and presents an engaging performance.	Stage presence and concert etiquette are somewhat professional and appropriate. Ensemble is not always focused on the conductor, and the performance is not always engaging.	Stage presence and concert etiquette are not professional or appropriate. Ensemble does not watch the conductor, and the performance is not engaging.

References

Abeles, H. (2016a). Alternative approaches to the assessment of the outcomes of musical

instruction.

Abeles, H. (2016b). Assessment strategies for the arts: online lectures.

Hafeli, M. (2009). What Happened To Authenticity? "Assessing Students' Progress And

Achievements In Art" Revisited. In R. Sabol and M. Manifold (Eds.), Through the Prism:

Looking Into the Spectrum of Writing by Enid Zimmerman. Reston, VA: National Art

Education Association

Wiggins, Grant (2010). Educative Assessment: Designing Assessments to Inform and

Improve Student Performance. San Francisco CA: Jossey-Bass.