The Effects of Using Self-Assessment Rubrics and Traditional Assessments on Student

Performance on Recorders in a Fourth-Grade Class

Caleb M. Shaw

Milligan College

Abstract

The purpose of this study was to examine the effects of using self-assessment rubrics and traditional assessments on 4th grade student performance on recorders. The sample consisted of fourteen 4th grade general music students – 8 females and 6 males. The sample was assigned to experimental and control groups. The experimental group was taught using skills-based self-assessment checklist while practicing to enhance skills associated with playing the soprano recorder. The control group was taught using traditional methods. Both groups were administered a pretest. Data were collected using a posttest after teaching both groups for 4 weeks. Data were analyzed using ANCOVA and multiple regression. The results revealed a significant difference between the experimental and control group ($F(1,13)=20.010$, $p=.001$) on the posttest. The second ANCOVA revealed no significant difference for the experimental group between male and female posttest scores when covarying the pretest score ($F(1,4)=6.175$, $p=.068$). The multiple regression equation to determine if average practice time or the number of days practiced had an influence on posttest scores was not significant ($F(3,3)=.332$, $p=>.05$), either. These results suggest self-assessment and skills-based practice help students master recorder skills.

Keywords: performance assessment, recorder, rubric, self-assessment, skill mastery

M|LLIGAN C⊕LLEGE

Academic excellence since 1866

Date: March 05, 2018
From: The Institutional Review Board (IRB) at Milligan College

Re: IRB Approval

Submission type: Initial Submission

Dear Caleb Shaw,

On behalf of the Milligan College Institutional Review Board (IRB), we are writing to inform you that your study *'The Effects of Using Self-Assessment Rubrics and Traditional Assessments on Student Performance on Recorders with Fourth-Grade Students'* has been approved as expedited. This approval also indicates that you have fulfilled the IRB requirements for Milligan College.

All research must be conducted in accordance with this approved submission, meaning that you will follow the research plan you have outlined here, use approved materials, and follow college policies.

Take special note of the following important aspects of your approval:

- Any changes made to your study require approval from the IRB Committee before they can be implemented as part of your study. Contact the IRB Committee at **IRB@milligan.edu** with your questions and/or proposed modifications.
- If there are any unanticipated problems or complaints from participants during your data collection, you must notify the Milligan College IRB Office within 24 hours of the data collection problem or complaint.

The Milligan College IRB Committee is pleased to congratulate you on the approval of your research proposal. Best wishes as you conduct your research! If you have any questions about your IRB Approval, please contact the IRB Office and copy your faculty advisor if appropriate on the communication.

Regards,
The IRB Committee

Table of Contents

Chapter 1

Introduction

Summative assessment has long been accompanied by rubrics. The use of rubrics to evaluate students is a thorough tool to help students become more aware of what exactly is being assessed and allow the teacher to assess individual strengths and weaknesses in both performance-based and written assessments. Rubrics allow the teacher to clearly measure skills and performance indicators while the student benefits from skill-specific feedback. This feedback also allows students to work on specific skills and improve in the areas most needed. In the context of music education, the teacher can assess using multiple types of rubrics and assess for both expression and technical aspects of a performance either jointly or separately (Valle, Andrade, Palma, & Hefferen, 2016).

Along with formal assessment in the classroom, rubrics serve a purpose outside of the classroom as self-assessment tools for students to use while practicing (Valle et al., 2016). Just as music teachers are trained to assess skills and performance indicators, students can be taught to recognize their own strengths and shortcomings in their own practice. Students too often are not taught effective practice techniques, and therefore their performance product suffers in the classroom and in concerts. Rubrics can be slightly modified to help students not only improve practice techniques, but also to provide feedback for specific skills. In order for this type of self-assessment rubric to effectively give feedback, students must be taught how to be critical of their own practice techniques and skill deficits. Teaching students to self-assess allows them to evaluate their practicing from a less personal viewpoint and allows them to connect their learning to the overall learning goals and standards being assessed (Hale & Green, 2009).

In order for any rubric, self-assessing or teacher administered, to be successful, the rubric must be created with explicit, succinct, limited, and easily understood criteria (DeLuca & Bolden, 2014). Students should know exactly what is being assessed and how the assessment will be measured. In order to increase their understanding, criteria should not be superfluous, but succinct, as well. The rubric should only assess skills that are necessary and should not include additional information or performance indicators. Rubrics must also be easy to decipher for students, as they need to understand what is expected of them.

Self-assessment rubrics include several key components that assist students in evaluating their practice at home (Valle et al., 2016). The first component is a list of basic skills. These skills are usually sequential and increase with difficulty as they are mastered in order. Along with skills, a simple response method is usually included for students to self-check their practicing and think critically about their work. Finally, teacher-created feedback is provided to accompany each skill. As students are unable to successfully complete certain skills, they fill out their rubric accordingly, and are then directed to appropriate feedback to help them improve the skill with which they are having trouble (Valle et al., 2016).

After students have practiced at home, self-assessed, and implemented feedback, the teacher can then create a rubric alongside the students to assess their performance on a summative assessment (Burrack, 2002). These summative assessments can assess singular performance indicators or many at a time, as well (DeLuca & Bolden, 2014). As previously stated, all assessments should at least be agreed upon by both teacher and students, if not jointly created (Burrack, 2002). The use of rubrics in many steps of assessment can help students develop better practice skills and become better performers both in and out of the music classroom. Self-assessment rubrics help students to think more critically about their rehearsal

and improve work ethic (Burrack, 2002) while traditional assessments utilizing rubrics in the classroom help students to understand what exactly is being assessed (DeLuca & Bolden, 2014).

Statement of the Problem

Students often times have to be explicitly taught to perform tasks that adults assume they have always known. One of these tasks for music students is effective practice techniques. Just as students are often unable to take notes effectively without being explicitly taught, they also must be taught good practice techniques for at-home practice. One method of practicing at home includes the use of rubrics to provide explicit feedback that helps students improve specific skills within a set needed to play recorder with correct technique (Valle et al., 2016). Students who are simply given a practice time log have no sort of accountability for the effectiveness of their practice time. Ideally, poor practice technique would be remedied with effective self-assessment and feedback. Therefore, the problem of this study was to determine if the use of self-assessment rubrics with explicit feedback would improve practicing and thereby improve performance on a traditional recorder performance assessment.

Purpose of the Study

The purpose of this study was to investigate the effects of the use of rubrics, self-assessment, and feedback from this self-assessment. Students were taught how to recognize and remedy improper playing technique through critical analysis of their own recorder playing. This self-assessment then either led the student to the next skill, if mastered, or to an explicit instruction describing how to help master that specific skill through better playing technique. Similar techniques were utilized by teachers in a past study with positive results (Valle et al., 2016).

Significance of the Study

It is crucial that music students develop good practice techniques at a young age in order to ensure success throughout their music education. Too often, music students become frustrated with practicing because they have not been taught good practice techniques. Teaching students to self-assess their practice could not only improve their practice skills, but also could help them to identify performance indicators that need more attention on their own rather than waiting for a music teacher to point out areas in need of improvement. The findings of this study will help to inform music educators of effective and ineffective practice techniques to teach their students involving self-assessment rubrics for instrumental practice.

Limitations

The following limitations were encountered in this study:

1. There was no way to guarantee the amount of practice time students dedicated to the use of self-assessment rubrics. Students were expected to tell the truth about their actual practice time and use of the rubrics.

2. The instruments were created by the researcher and were not tested for validity or reliability.

Definitions

1. **Fingering** – The proper placement of fingers on holes when playing recorder

2. **Performance indicator** – Evidence of a skill being executed while playing recorder

3. **Self-assessment** – The process of critically examining one's own work or, in this case, practice techniques

4. **Self-assessment rubric** – A self-assessment instrument that utilizes a list of skills, some form of self-assessment response, and explicit instructions on how to fix certain deficits

5. **Tone** – The quality of sound produced when playing an instrument or singing

6. **Traditional Assessment** – Assessment administered and adjudicated by the teacher that identified correctness of performance through the use of performance and technique indicators. Students performance is measured on a rubric point system to identify correctness of performance technique.

Overview of the Study

This study is broken down into five chapters. This first chapter contains an introduction to the study, statement of the problem, purpose, definitions, and the significance of the study. The second is a review of pertinent literature which relates to the topic. The third chapter includes the methodology, procedures, research questions, and all information regarding the collection of data. Chapter four contains the results of the study. Chapter five, the final chapter, includes a review of the study as well as conclusions, recommendations for further research, and implications.

Chapter 2

Review of the Literature

Introduction

Effective instrumental practice techniques allow students to practice more efficiently and help them to improve at specific skills when practicing. However, effective practice techniques can vary from student to student, with some methods working better for some than others. However, research suggests that students can adopt practice methods such as rubrics that utilize performance indicators, explicit feedback, and skills checklists in order to better their practice technique (Valle, Andrade, Palma, & Hefferen, 2016). Practicing using traditional methods like practice logs and theory worksheets simply don't produce the required results needed to succeed and excel in playing an instrument at any level. By simply going through the motions of practicing, students don't typically take the time to reflect on what they have done well and what they need to improve. A qualitative study by Oare (2016) sought to determine effective practice techniques for early instrumentalists. In triangulating student practice techniques, verbal cues when practicing, retrospective interviews, and results of practice, he determined that students' practice is generally different for each one. He determined that thinking aloud and developing an aural image of the music during practice could lead to more motivation to continue and thereby to higher achievement on skill-based performance indicators.

In spite of differences in learning and practice, the use of self-assessment rubrics can help students evaluate their playing and can offer helpful feedback through a series of skill checks with explicit feedback (Valle et al., 2016). The primary research questions being posed for this review are: Do self-assessments that are utilized to give both student and teacher formative feedback help students to achieve higher results on traditional performance assessments? Also,

which types of rubrics, practice methods, and traditional assessment methods are best for evaluating instrumental performance?

<div align="center">Rubrics for Traditional Assessment</div>

In order to answer these questions, rubrics must be explored in multiple scenarios for grading. One source defines a rubric as "a visual narrative of the criteria that defines and describes the important components of an assignment" (Balch, Blanck, & Balch, 2016, p. 20). Evaluative rubrics are useful both for the student while practicing at home as well as in whole-group rehearsal for the teacher. Teachers often make mental checklists of their expectations during a rehearsal and guide each rehearsal towards attempting to check each component off of this list. However, it is helpful to communicate these expectations to students explicitly so that they know for what they should strive. In fact, students should have an active role in forming these expectations along with the teacher in order to increase understanding (Hale & Green, 2009). Teachers must be certain that students understand expectations and the performance indicators being assessed in order for them to be successful using self-assessment and traditional assessment rubrics. Also, teachers must be sure to use as simple or complex of criteria as appropriate for an assessment (Hale & Green, 2009). Unless multiple higher order thinking criteria are being assessed at once, rubrics should be made as simple as possible for students to best understand them (DeLuca & Bolden, 2014). Traditional assessment rubrics should be clear, describe behaviors that will make up a performance, and give varying degrees of correctness regarding these behaviors (Balch, Blanck, & Balch, 2016).

In a general music classroom, whole-group instruction is a very typical means of teaching. Therefore, assessment must be tailored to the short time frame of a general music class and should be done as quickly and often as possible (Dunbar, 2012). This type of whole group

assessment could manifest as checklists or other types of concise rubrics used by the teacher. Most grading in secondary performance-based music classes is based largely on attendance, behavior, and participation, rather than achievement (Pellegrino, Conway, & Russell, 2015). This poses a problem for music teachers, as they are often unaware of the achievement level of each individual student. An assessment method other than those listed above must be implemented in order to ensure that each student is learning in whole-group instruction. Any individualized assessment tool to be used during whole-group instruction must be "usable, accurate, and quick" (Dunbar, 2012, p. 4) in order for it to be effective in short class periods. This individualized whole-group assessment should mimic traditional assessment as well (Dunbar, 2012).

In one action research study conducted by Hood (2012), students were given an active role in selecting criteria on which they would be evaluated. After the first implementation of this criteria selection involvement, data were analyzed from assessments and methods of criteria selection and student involvement were adapted according to Hood's findings. Throughout three cycles of this action research, it was found that involving students in the criteria selection process significantly improved assessment grades and increased their self-confidence and self-efficacy (Hood, 2012). This study shows that the student's thorough knowledge of assessment criteria can improve achievement through increasing comfort level with the written or performance material being assessed.

Teachers can use performance rubrics every day in class if desired, but they are especially useful for both teacher and student evaluation of performances as summative assessments. These types of rubrics can help to assess both expressive and technical aspects of a performance either separately or in tandem. It is good to note that teachers must begin any unit or lesson with the end goal in mind. Students should be made aware of what will be expected of

them at every step of learning (Boyd, 2013). For younger students, it is perhaps best to assess individual skills rather than multiple, so as not to confuse them. However, as students' abilities increase in difficulty, multiple skills can be assessed at once. Using a discrete-component rubric would be ideal for younger students, as it assesses only one skill at a time (DeLuca & Bolden, 2014). Students who have a firm grasp on the skills being assessed, however, can be assessed using an integrated-component rubric. This type of rubric connects performance indicators together in order to assess higher thinking and more in depth performance (DeLuca & Bolden, 2014). Also, it is just as important to evaluate the overall performance as it is to evaluate each separate aspect. Music is, after all, an art that should be appreciated for the sake of art itself.

Rubrics for summative assessment are effective at helping students to understand exactly what is being assessed, so long as students are privy to the expectations and criteria on a rubric. However, it is important that students learn skills and terminology that will be included on the rubric as skill learning is initially taking place (Hale & Green, 2009). When planning for instruction and assessment together, rubrics connect to instruction by helping the teacher predetermine which standards are most important and what will actually be assessed at the end of the lesson or unit (Cooper & Gargan, 2009). This discernment allows for teachers to focus on the most important factors of a lesson or unit when teaching, as well.

Rubrics are not a perfect method for summatively evaluating students, however. While rubrics allow for more diverse authentic assessments, they have several downfalls (Cooper & Gargan, 2009). For instance, grading using rubrics is often very subjective to the assessing teacher. They can be burdensome and time consuming to use for both students and teachers alike. Rubrics can also be restrictive for more creative students who like to think "outside the box" (Cooper & Gargan, 2009, p. 7). The listing of concise, explicit criteria often does not allow for

any variance on the part of the student, even when variability could demonstrate higher-order thinking skills.

Traditional performance assessments using rubrics can also be unreliable depending on the conditions in which they are used. In a study by Martin (2007) involving ten certified adjudicators and eight high school wind performers, research showed that at least five adjudicator ratings were necessary in order to create a reliable assessment. A one-adjudicator, one-occasion performance produced the least reliable evaluation results of all. This study also showed that the number of times that a student performed had little effect on their scores. Because so many music educators are one of a kind in their schools, a lack of adjudicators could potentially produce reliability and validity concerns on a day-to-day basis when evaluating student performance. These findings indicate that a reliable measurement of performance is nearly impossible for the typical music teacher because of subjectivity and preference. This subjectivity leads to a need for better rubrics with more stringent criteria, performance indicators, and rating scales (Martin, 2007).

In a study to find the effect of weighted rubrics on choral and instrumental adjudication, the effect of a new weighted rubric on adjudicating Kansas state high-school large group festivals was analyzed. It was found that universal weighted rubrics, which evaluated performance and placed more weight on skills that required critical thinking, were internally consistent with moderately high results (Latimer, Bergee, & Cohen, 2010). Weighted rubrics were also found to have moderately high reliability when compared to other researched methods of adjudicating large music groups, including choral and instrumental.

However, Saunders and Holahan (1997) conducted a study involving 926 students auditioning for Connecticut All-State Band woodwinds and brass sections that had no weight

placed on criteria. The adjudicators utilized a criteria-specific rating scale to evaluate solo

performances of each student. Thirty-six judges were given the same criteria-based rating scale

with which to judge. Upon completion and analyzation of the data from the score sheets, it was

found that this evaluation method yielded substantially reliable results from adjudicators. The

use of a gradually increasing rating scale rubric in evaluation allows teachers to decrease

subjectivity when grading by displaying specific performance indicators (Saunders & Holahan,

1997).

These two studies were in agreement in regards to the need for specific requirements in

performance assessment rubrics. In order to create less subjective rubrics, evaluators must

incorporate skill-specific, concise, explicit criteria that students can easily rehearse with and can

be easily evaluated as unbiased as possible (Cooper & Gargan, 2009; DeLuca & Bolden, 2014).

However, using only summative assessments as a means of assessing student performance does

very little toward advancing students' musically (Wesolowski, 2012). When students utilize self-

assessment as formative assessment, they are better able to correct improper technique before it

becomes a habit.

<div align="center">Self-Assessment Rubrics</div>

One of the most effective ways to help students succeed with evaluations using rubrics is

to teach them to self-assess using rubrics that include the same criteria on which they will be

evaluated (Valle et al., 2016). Andrade (2010) defines self-assessment as "a process of formative

assessment during which students reflect on the quality of their work, judge the degree to which

it reflects explicitly stated goals or criteria, and revise accordingly" (p. 3). The idea that self-

assessment must be formative suggests that it also much be task-specific (Andrade, 2010). Self-

assessment would serve no purpose were the student unable to learn from and remedy skill

deficits. Andrade (2010) also suggests that students must be self-regulated learners in order to employ successfully these self-assessment strategies.

Students can benefit greatly from self-assessing their performance in order to grow as a musician. This cyclical learning process can help students recognize skills they have mastered as well as skills they need to work on (DeLuca & Bolden, 2014). Self-assessments of performances can also help students to improve their work ethic and overall performance in the future (Burrack, 2002). Helping students to self-assess helps them to see their work from a less personal viewpoint, as though they were evaluating themselves from someone else's position (Hale & Green, 2009). This helps students to see the value of standards-based learning. When students put themselves in an evaluator's position, they tend to see their own flaws more clearly and are therefore better able to think critically about what can be done to help build skill level (Hale & Green, 2009). These performance self-assessments can be completed either by simply reflecting on a performance or practice session, or by watching or listening to a recording of the performance or practice session. Students who are asked to do this, would need to be well aware of the skill expectations as well as the performance indicators on which to assess themselves.

Just as teachers are taught that students who make the rules are more likely to follow them, students who play an active role in developing criteria for assessment rubrics are more likely to understand and succeed at meeting expectations (Burrack, 2002). It is much more likely that all students will better understand what is expected when creating the criteria of a rubric for themselves than if the criteria is exclusively decided by the teacher. This method of creating a rubric can also give the teacher a better idea of what students view as important when evaluating their playing or singing (Burrack, 2002). In some cases, students might even develop a more difficult assessment for themselves than the teacher would. This manner of creating criteria for

assessment allows students to express their own level of expectations for themselves, which helps the teacher to scaffold instruction to be more rigorous for some (Boyd, 2013). The teacher does need to be conscious of keeping student-centered criteria as explicit as possible. It is essential that each performance indicator is concise and explained as well as possible to encourage student understanding and success (Burrack, 2002; DeLuca & Bolden, 2014).

Another facet of self-assessment is to use rubrics to help inform individual practice techniques after the regular school day has ended. In order for these practice checklists to work as intended, students must have a firm grasp on the types of skills that are being assessed and be able to independently correct their practice methods to advance toward mastery of skills. (Hale & Green, 2009). If students are unable to comprehend what it is that needs improvement, then they will be unable to master skills independently without face-to-face teacher-given feedback (Hale & Green, 2009).

Andrade (2007) suggests that self-assessment procedures should be accomplished in three steps. The first of these involves the teacher providing a rubric that includes desired performance expectations and possibly some sort of modeling example. In the second step, students create a sort of draft of their work and either formally or informally compare their work to the rubric. Finally, students can use their self-assessment findings from the rubric to make corrective adjustments to their work. This process of self-assessment helps students to understand what they have done correctly as well as incorrectly, and helps them to accurately fix problems they are having. When paired with specific skill instruction, this process can help students master these skills in progressive order. This self-assessment also helps students to see their work from a less personal viewpoint (Hale & Green, 2009).

The use of self-assessment forms or rubrics can help students to understand their deficits and remedy them. In a study by Yarbrough (1997), higher education choral conducting students were asked to learn basic conducting skills through learning a series of scaffolded skills. After each skill was rehearsed, students would watch a video of themselves conducting those skills and complete a self-evaluation form. After their self-assessment, they were assessed by the instructor in a posttest. This continued until all skills were accomplished. This study yielded significantly positive results between accurate self-evaluation and posttest scores (Yarbrough, 1997). While students were learning to self-assess from videos of their performances, they were also learning to critically evaluate their rehearsal techniques through the scaffolding of skills being taught. This technique is effective utilizing videos of the performances and practices, but students could possibly replicate this type of practice in an at-home rehearsal setting.

In order for self-assessment to be effective, explicit feedback must be given in some way to help students fix problems they are having with specific skills (Wheeler, 2016). In a study by Wheeler (2016), two groups of middle school students were given verbal feedback (treatment) and no verbal feedback (control) to find the effect of explicit verbal feedback on self-assessment of recorder performances. While the experimental group began with worse results, their self-assessment ability soon improved more than the control group. This suggests that skill-based explicit feedback given from teachers can have a positive effect on students' ability to self-assess. It was also found that beginning ability level had no effect on skill growth when self-assessing - all levels of performers benefited from explicit feedback.

A mixed-methods study by McKevitt (2016) compared two groups of students to find the effect of direct feedback on the writing process. The comparison group were given a self-assessment rubric with which to compare their rough drafts of an essay. The experimental group

were given specific feedback regarding their drafts from a tutor. It was found that students who received direct feedback throughout the revision process from a tutor benefited more than students who only compared their rough drafts to a rubric. This translates to music in the form of explicit written or verbal feedback concerning practice techniques and performance as well.

Success in utilizing self-evaluation to improve practice methods would involve teachers first articulating their expectations explicitly and concisely (Valle, Andrade, Palma, & Hefferen, 2016). Communicating expectations and performance indicators should be accomplished in class before the students are expected to utilize the practice checklist at home. Along with these expectations, students should also know several tips that can help them improve at skills with which they are struggling (Valle et al., 2016). This prescribed feedback should be explicit as well, and should be unique for each skill being attempted. This feedback should be well-written as to allow for mastery of each attached skill, as well. For example, if students are unable to successfully play the rhythm of a piece of music, the feedback instructions could read "Tap the rhythm using syllables and see if it matches what you played on recorder." This type of feedback is directly connected to each skill being attempted and utilizes the discrete-component type of rubric in self-assessment (DeLuca & Bolden, 2014). Students must also be able to evaluate themselves based on the expectations set forth by the teacher and themselves (Valle et al., 2016). Using rubrics and checklists for corrective feedback cannot be an effective method of practicing without students' ability to self-diagnose problems in their skillset. In order for a student to fix a detriment, they must first recognize that there is a problem to fix and understand how to utilize feedback on the rubric for their benefit.

The last step in utilizing rubrics to benefit practicing at home is for students to revise their practicing to reflect the corrective feedback that the rubric provides (Valle et al., 2016). In

order to revise correctly, the skills and corrective feedback should be listed sequentially and the students should follow the chain of corrections so that each skill builds on past mastery. For example, students should not move on to attempting to master playing more difficult fingerings on recorder before mastering the more rudimentary positions. This same method of self-assessment can be applied to whole-group performance self-assessment and peer-to-peer assessment as well. If students can build their mastery, skill by skill, they can gain independence and will achieve their goals of mastering all the skills and performance indicators included in the rubric. After students demonstrate their learning by filling in the rubric, one skill at a time, the teacher can then assess how well the rubric worked in relation to practicing for mastery of certain skills.

A downfall of self-assessment is that students sometimes tend to rate themselves higher on self-assessments than an expert or teacher would. A study by Hewitt (2002) found that junior high instrumentalists' self-evaluation scores generally rose over time. However, self-evaluation accuracy did not rise along with these scores. When compared to expert evaluations of the same performances over time, students consistently rated themselves higher than evaluators. This suggests that students overestimate their achievement when compared to experts' evaluations. It is of utmost importance to teach students realistic expectations when practicing to ensure that they don't overestimate their performance and practice less.

While self-evaluation is important for the process of learning, Andrade (2007) insists that there is a staunch difference between self-evaluation, which involves summative self-grading, and self-assessment. Students utilizing self-assessment engage in "a process in which [they] collect information about their own performance and see how it matches their goals and/or the criteria for their work" (Andrade, 2007, p. 160). This definition insists that self-assessment

should be used for advancement of skills toward mastery rather than for grading purposes. Therefore, students should view self-assessment less as an obligatory grading procedure, and more as a tool to help themselves advance toward a goal more effectively.

It is crucial that self-assessment, self-evaluation, and assessment by the teacher all align with learning goals, having matching criteria, and progress students toward mastery of necessary skills (Andrade, 2010; Burrack, 2002; Cooper & Gargan, 2009). With these aspects aligned, students are more likely to succeed and push toward mastery at a faster pace.

Self-Efficacy

While self-assessment and advancement of skill sets plays a large role in student success in playing an instrument, self-efficacy and confidence are also a large part of ensuring student success. Research indicates that students who feel they can conquer a skill or content area are more likely to succeed when attempting to master it. This can also lead to better intrinsic motivation to learn and build on previously mastered skills (Schunk, 1991). Self-assessments as formative assessments help students to stay on track with their learning. However, the degree of accuracy of these self-assessments is important for students to succeed at performance and academic success (Chen & Zimmerman, 2007).

Teachers must directly involve students in the learning process to ensure their success. Teaching learning strategies and instilling positive beliefs can help to raise self-efficacy for some students, and thereby can raise achievement in performance and mastery of skills (Hewitt, 2015). In a study conducted by McCormick and McPherson (2003) involving 332 instrumentalists between the ages of 9 and 18, it was found that self-efficacy prior to an assessment was the best predictor of a musician's performance. This study found that if students believe that they won't be able to play a piece of music correctly the first time, they are more likely not to attempt it.

Teachers must be sure that student understand how to practice and that they have confidence in themselves so that they believe that they can succeed. One method to help with this is to give early feedback that highlights success, as this can lead to higher self-efficacy and more effective learning from the start (Schunk, 1991).

Self-efficacy also strongly connects to self-assessment and evaluation of performances. In a study involving 354 high school students, participants were placed into two groups to discover the relationship between self-efficacy prior to a performance and students' perceived self-evaluations post-performance (Hewitt, 2015). This study found a strong and positive correlation between these two measures. Performance evaluations, however, did not necessarily explain these data. While there are some concerns with these data regarding accuracy of student self-evaluation scoring, the correlation between self-efficacy and perceived performance achievement shows that students who believe in themselves and are confident have a better feeling about their performances and are more highly evaluated.

<div align="center">Conclusion</div>

Overall, most studies have returned positive results in favor of utilizing self-assessments in order to inform students of progress when practicing (Andrade, 2010; Yarbrough, 1987; Hale & Green, 2009). There are also favorable results in regard to using the same criteria on summative assessments as what students are used to on formative self-assessments. This familiarity plays a huge role in students' confidence, and thereby their performance. Formulating a reliable, valid, unbiased rubric on which to evaluate students is the main challenge for sole music teachers in their schools. The best way to accomplish this is to allow students to be a part of the criteria selection process, to make all expectations clear to students, and finally be sure that students utilize similar rubrics during self-assessment as formative assessment to help them

master skills in a sequential manner (Andrade & Du, 2007; Burrack, 2002; DeLuca & Bolden,

2014).

Chapter 3

Methodology and Procedures

At-home practice methods are often ineffective at helping students master skills to help mimic classroom practice and teaching. The purpose of this study was to determine if using rubrics and checklists to self-assess can help students to practice more efficiently and master skills sequentially through at-home practice as compared to using only traditional assessments with no change in study habits. Based on the review of literature, there is evidence that the use of self-assessment as formative assessment using rubrics and/or checklists may help students achieve mastery in a sequential manner through informed practice. A thorough understanding of the skills being mastered promotes a high self-efficacy, which tends to boost achievement.

This chapter includes information about the population of the school in which the study was conducted, as well as participant selection procedures. It also includes methods, procedures, and descriptions of instruments and data collection procedures.

Population

The population for this study was comprised of elementary students at a rural Northeast Tennessee public school. The school was made up of 449 students, most of whom were white. 94.2% of the students were white, 3.1% were Hispanic of Latino, 2.2% were African American, and less than 1% were Asian. Approximately 40% of students were economically disadvantaged and 17.6% were students will disabilities.

Sample

The sample of students for this research were from two fourth-grade classes who attended general music class once per week. Both of the classes had roughly twenty-five students. Students were randomly selected to form experimental and comparison groups. All students in

both groups were taught using traditional beginner instrumental teaching methods, with the comparison group receiving an additional treatment. Both groups performance was compared after they were taught using traditional methods and the comparison group was given the treatment.

Data Collection Instruments

Three instruments were used to collect data throughout this study. The first instrument was a pre-test given to both control and experimental groups before the treatment was applied to the experimental group. The pre-test was a performance-based assessment given to each participant on an individual basis. The students were given a short piece of music to practice on recorder for one week with no additional instruction or tools to use. The students then played the piece for the teacher and were rated using a rubric for the following skills: pitch/fingering, rhythm accuracy, and breath control. Each teacher assessment was videoed on an iPhone using a tripod and the video was reviewed to ensure accurate scoring. This video captured the sound and visual cues that informed scoring using the rubric.

The second instrument utilized was a checklist for students in the experimental group to take to help them practice at home. The checklist included a skills column that listed recorder playing skills in sequential order to help students master one skill at a time and build on the past skills learned. The skills included: correct fingering, correct pitch played, rhythm accuracy, and breath control. There were also "yes" and "not yet" columns for students to self-assess their playing as they practiced. To help students master skills, explicit feedback advice was also given beside each skill to help students fix problem areas that they had trouble with by providing helpful tips for correcting mistakes associated with each skill. For example, next to the "correct

fingering" skill, the feedback stated, "Double check your fingering with the note being played on your fingering chart."

The final instrument used was identical to the first, with the addition of a short survey to determine for how many days and for how long each day the student rehearsed at home. Students were filmed on an iPhone for these assessments similar to the pre-test. The data collected from these instruments and assessments were statistically analyzed following their implementation. The pre-assessment and summative assessment rubrics consisted of varying levels of mastery for each skill – no mastery was worth no points, some mastery was one, approaching mastery was two, and mastery was three. This quantification of data helped with analyzation. It should be noted that none of these instruments were checked for validity or reliability prior to the study. The control group was given instruments one and three, but were not given the treatment instrument.

Procedures

Prior to the start of this study, permission was obtained from the system and school in which the research was taking place, as well as from the Milligan College IRB. Also, consent forms were signed and returned from all parents and guardians for the participants in the study. Parents and students were informed of and assured of confidentiality. They were informed of their right to withdraw from the study at any time, as well. They were also assured that no mental or physical harm would come to anyone involved, except for the possible mild discomfort or stress associated with performance-based assessments.

Prior to the study, students in the fourth grade were already assigned into their two respective classrooms, therefore no random assignment was possible. Students in both of these classrooms were given identical general music and recorder instruction prior to the start of the

study. Students had instruction in recorder in third grade, and participated in several review lessons prior to the study, as well. The purpose of the study was not to determine how well students retained new information, but rather, how well they could sequentially build skills independently through practice of skills already taught. Therefore, the students only utilized the four pitches that were taught the year before: G, A, B, and C.

At the start of the study, seven students were selected for the control and treatment groups. These fourteen students were given a pre-assessment to determine their mastery level for the skills being assessed while playing a short teacher-made recorder piece. Next, the control group of students were given skill-focused instruction similar to the prior weeks during class and asked to practice ten minutes per night each weeknight until the next week. The experimental group was given skill-focused instruction like the control group, but were also given the skills checklist to take home. They were instructed on how to utilize the checklist and explicit feedback to help self-assess their practicing. They were taught to use the instrument to sequentially build on skills. They were then asked to return these filled out the following week.

A summative assessment was then given to all participants to determine if utilizing the checklist and explicit feedback during at-home practice helped students master skills necessary for successfully playing recorder. A short survey was also given to determine the amount of practice each day and the number of days practiced. Data were collected from these assessment instruments, quantified, and statistically analyzed to determine the practice checklist's effectiveness.

Research Question and Hypotheses

Research Question 1: Is there a difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score?

Research Hypothesis 1: There is a difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score.

Null Hypothesis 1: There is no difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score.

Research Question 2: Is there a difference between genders' scores for the treatment group on the posttest when covarying the pretest scores?

Research Hypothesis 2: There is a difference between genders' scores for the treatment group on the posttest when covarying the pretest scores.

Null Hypothesis 2: There is no difference between genders' scores for the treatment group on the posttest when covarying the pretest scores.

Research Question 3: What is the impact of the number of days practiced and average practice time on posttest scores?

Research Hypothesis 3: The number of days practiced and average practice time have an impact on posttest scores.

Null Hypothesis 3: The number of days practiced and average practice time do not have an impact on posttest scores.

Chapter 4

Data Analysis

The purpose of this study was to examine the effects of using self-assessment rubrics and traditional assessments on student performance on recorders with fourth-grade students.

Data Collection

Data were collected from a sample of fourteen music students in the fourth grade. The participants were assigned into two groups: the control group, which received traditional recorder instruction and no additional at-home practice methods, and the treatment group, which received traditional recorder instruction as well as instruction on how to use skills-based practice checklists. The demographics of the participants are displayed in table 1. Students were given recorder instruction for three weeks prior to the initial pretest, and were playing the pitches G, A, B, and C during this instruction. Data were collected for both a pretest and a posttest following these weeks of instruction to determine the effectiveness of using checklists when practicing. Following these tests, students were given a short survey to determine the number of days spent practicing each of the two weeks, as well as the average time spent each day. A pretest was given to both control and treatment groups prior to the treatment being applied.

Table 1

Demographic Profile for the Participants

Group	Frequency	Percent %	Number of Females	Number of Males
Control	7	50.00	3	4
Treatment	7	50.00	5	2
Total	14	100.00	8	6

Research Questions and Related Hypotheses

Three research questions were formulated to determine the effects and potential benefits of using self-assessment rubrics and traditional assessments on student performance on recorders. **Research Question 1:** Is there a difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score?

In response to research question 1, an analysis of covariance test was calculated to examine the influence of the treatment on posttest score covarying out the effects of pretest score. **Research Hypothesis 1:** There is a difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score.

A one way between-subjects ANCOVA was calculated to examine the influence of the treatment on posttest score covarying out the effects of pretest score. The pretest score was significantly related to the posttest score ($F(1, 13)=20.010$, $p=.001$). The main effect for grouping (control and treatment) was significant as well ($F(1,13)=11.509$, $p<.05$) with the treatment group performing significantly higher ($M=4.57$, $sd=2.070$) on the posttest than the control group ($M=1.86$, $sd=2.478$). The Eta^2 was 0.511, which means that 51.1% of the variance for the posttest could be explained by the pretest score. 48.9% of the variance could be explained by other variables. Therefore, the null hypothesis was rejected. A summary for the ANCOVA test can be found below in table 2.

Table 2

ANCOVA Summary for the Posttest

Source	SS	df	MS	F	p	Eta^2
Pretest	40.376	1	40.376	20.010	.001	.645
Group	23.222	1	23.222	11.509	.006	.511
Error	22.195	11	2.018			
Total	233.000	14				

Research Question 2: Is there a difference between genders' scores for the treatment group on the posttest when covarying the pretest scores?

In response to research question 2, a one way between-subjects ANCOVA test was calculated to compare the effect of gender on the posttest score for the treatment group when covarying the pretest scores.

The following research hypothesis was related to research question 2:

Research Hypothesis 2: There is a difference between genders' scores for the treatment group on the posttest when covarying the pretest scores.

A one way between-subjects ANCOVA was calculated to examine the influence of gender on the posttest score covarying the effects of the pretest score. The pretest score was not significantly related to the posttest score ($F(1,4)=6.175$, $p=.068$). The main effect for gender (male and female) was not significantly related to the posttest score either ($F(1,4)=1.969$, $p=.233$). The mean for males was ($M=4.50$, $sd=2.121$) and the mean for females was ($M=4.60$, $sd=2.302$). The Eta^2 for the main effect was .330, which means that 33% of the variance for the posttest could be explained by the gender. 67% of the variance could be explained by other

variables. Therefore, the null hypothesis was retained. A summary for the ANCOVA test can be found below in table 3.

Table 3

ANCOVA Summary for the Genders' Posttest Scores

Source	SS	df	MS	F	P	Eta^2
Pretest	15.597	1	15.597	6.175	.068	.607
Gender	4.974	1	4.974	1.969	.233	.330
Error	10.103	4	2.526			
Total	172.000	7				

Research Question 3: What is the impact of the number of days practiced and average practice time on posttest scores?

Research Hypothesis 3: The number of days practiced and average practice time have an impact on posttest scores.

In response to research question 3, a multiple linear regression analysis was calculated to predict posttest score for the treatment group based on the number of days practiced and average practice time. The regression equation was not significant ($F(3,3) = .344, p>.05$) with an R^2 of .256. This suggests that 25.6% of the variance in posttest score was explained by the predictor variables. Results also suggest that 74.4% of the variance could be explained by other variables other than the predictor variables. 1st week practice time had a beta score of .628 ($p=.555$), 2nd week practice time had a beta score of -.495 ($p=.477$), and average time spent practicing had a Beta score of -.118 ($p=.897$). Neither the number of days or the average amount of time spent practicing was a significant predictor of the posttest score using a performance based assessment.

Therefore, the null hypothesis was retained. The multiple regression analysis results can be found in table 4.

Table 4

Multiple Regression Coefficients Summary

Variable	Beta	t	Sig.
1st Week Practice Time	.628	.662	.555
2nd Week Practice Time	-.495	-.811	.477
Avg Time Spent Practicing	-.118	-.140	.897

Chapter 5

This chapter includes a summary of findings, conclusions, recommendations, and implications of the research conducted to determine the effects of using self-assessment rubrics and traditional assessments on student performance on recorders with fourth-grade students.

Summary of Findings

When the first research question "Is there a difference in students' performance on the posttest when they are taught using self-assessment rubrics and traditional practice methods when covarying the pretest score," was examined, the results suggested that there was a significant difference in posttest scores for the control and treatment groups. Scores for the treatment group were significantly better when students were taught to self-assess and utilize a practice checklist rubric during at-home practice. Therefore, the null hypothesis was rejected. The primary treatment for this study was instruction on how to utilize a skills-based practice checklist and self-assess practicing at home. The experimental group students likely performed better on the posttest because of this self-assessment instruction and the explicit feedback given on the checklist. According to the literature, self-assessment rubrics with explicit feedback help students stay focused and allows them to master skills through self-efficacy (Burrack, 2002; Schunk, 1991). The inclusion of explicit feedback to accompany skills-based learning helped students to thoroughly self-assess, and thereby helped them master the skills being taught (Wheeler, 2016).

In regards to the second research question "Is there a difference between genders' scores for the treatment group on the posttest when covarying the pretest scores," it was found that there was no significant difference in male and female posttest scores ($F(1,12) = 3.516$, $p > .05$) when a

one way between-subjects ANCOVA test was calculated and the pretest score was covaried. This test was performed for the treatment group, with a total of two males and eight females. Female students had a mean score of (M=4.60, sd=2.302) and males had a mean score of (M=4.50, sd=2.121). There was no significant difference between the male and female scores on the posttest (F(1,4)=6.175, p=.068). This indicates that males and females received similar scores when covarying the pretest score. Therefore, the null hypothesis was retained. The results suggest that when students are taught using self-assessment checklists, male and female students benefit equally.

The final research question, "What is the impact of the number of days practiced and average practice time on posttest scores," was examined and a multiple linear regression analysis was calculated to predict posttest score for the treatment group based on the number of days practiced and average practice time. The regression equation did not show that practice time or number of days practiced had significant influence on the posttest score. The R^2 was .256, which means that only 25.6% of the variance could be explained by the predictor variables. The results for this question were likely influenced by a lack of practice time throughout the study. When given a survey to list practice times, eight of the fourteen students admitted to having only practiced 0-1 days each week, while the remaining six practiced between 2-5 days each week. In addition, no student practiced more than ten minutes each night that they practiced. These data indicate that sporadic practice habits likely influenced the effectiveness of the multiple regression test. While effective practice methods were taught for both groups, a lack of at-home practice time and effort likely influenced the results.

Conclusions

The purpose of this study was to determine the effects of using self-assessment rubrics and traditional assessments on student performance on recorders with fourth-grade students. Results of the various tests run indicated that the use of self-assessment and traditional assessments improved posttest scores when coupled with explicit feedback and skills-based practice methods with fourth-grade recorder students. However, the results also indicated that males and females in the treatment group did not achieve significantly different scores on the posttest when covarying the pretest score. In addition, students' scores were not influenced by the amount of time spent practicing for the performance-based assessment or the number of days practiced. Therefore, the research hypothesis was only proven for the first of the three research questions.

Recommendations

1. Further research should include a larger sample. This study only included seven participants in each group. A larger sample size would improve accuracy and reliability of the results of the study.

2. Further research should require that all students document and turn in their self-assessment checklists. Data should be analyzed for this instrument to test the effectiveness of the rubric itself in conjunction with the performance based assessment results. This could provide the researcher with an idea of the influence of the rubric itself on the pretest and posttest scores.

3. Validity and reliability testing should be performed for the self-assessment checklist to ensure that the instrument effectively teaches skills mastery.

4. Further research should require that students document their practice time each night as it happens, rather than collecting these data altogether at the end with a survey instrument. This would provide a more accurate representation of practice amount than the final survey.

Implications

Based on the results of the study:

1. Teachers should recognize the importance of skill-based teaching for elementary students in the general music setting. Teaching students to master basic skills in a scaffolded manner can help them succeed at mastery of an activity as a whole.

2. Parents should stress the importance of focused at-home practice time for instrumental students. They must play an active role in their child's practice to ensure that students are practicing the necessary amount as well as in a manner that teaches them to self-assess and learn from their mistakes.

3. Students must engage in rigorous, scaffolded skills-based practice to maximize their results. Students should build on skills that they have already mastered to be successful at playing an instrument at a whole.

References

Andrade, H. (2010). Proceedings from NERA Conference 2010: Students as the definitive source

of formative assessment: Academic self-assessment and the self-regulation of learning.

Storrs, CT. http://digitalcommons.uconn.edu/nera_2010/25

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment.

Assessment & Evaluation in Higher Education, *32*(2), 159-181. doi:

10.1080/02602930600801928

Balch, D., Blanck, R., & Balch, D. H. (2016). Rubrics— Sharing the rules of the game. *Journal*

of Instructional Research, *5*, 519-549. Retrieved from https://eric-ed-

gov.milligan.idm.oclc.org/?id=EJ1127629

Boyd, J. (2013). Idea bank: Progress through incentives: How one music program helps students

progress to higher levels of musicianship. *Music Educators Journal, 100*(1), 23-25.

Retrieved from http://www.jstor.org.milligan.idm.oclc.org/stable/43288762

Burrack, F. (2002). Enhanced assessment in instrumental programs. *Music Educators Journal,*

88(6), 27-32. Retrieved from http://www.jstor.org.milligan.idm.oclc.org/stable/3399802

Chen, P., & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-

efficacy beliefs of middle-school mathematics students. *The Journal of Experimental*

Education, 75(3), 221-244. doi:10.3200/JEXE.75.3.221-244

Cooper, B. & Gargan, A. (2009). Rubrics in education: Old term, new meanings. *The Phi Delta*

Kappan, *89*(1), 54. Retrieved from

http://www.jstor.org.milligan.idm.oclc.org/stable/42926555

DeLuca, C., & Bolden, B. (2014). Music performance assessment: Exploring three approaches

    for quality rubric construction. *Music Educators Journal, 101*(1), 70-76. doi:

    10.1177/0027432114540336

Dunbar, L. (2012). Performance assessment of the masses in 30 seconds or Less. *General Music*

    Today, 25(2), 31-35. doi: 10.1177/1048371311406266

Hale, C. L. & Green, S. K. (2009). Six key principles for music assessment. *Music Educators*

    Journal, 95(4), 27-31. doi:10.1177/0027432109334772

Hewitt, M. P. (2002). Self-evaluation tendencies of junior high instrumentalists. *Journal of*

    Research in Music Education, 50(3), 215-226. Retrieved from

    http://www.jstor.org.milligan.idm.oclc.org/stable/3345799

Hewitt, M. P. (2015). Self-efficacy, self-evaluation, and music performance of secondary-level

    band students. *Journal of Research in Music Education, 63*(3), 298-313.

    doi:10.1177/0022429415595611

Hood, A. A. (2012). Whose responsibility is it? Encouraging student engagement in the learning

    process. *Music Education Research, 14*(4), 457-478. doi:10.1080/14613808.2012.703174

Latimer Jr., M. E., Bergee, M. J., & Cohen, M. L. (2010). Reliability and perceived pedagogical

    utility of a weighted music performance assessment rubric. *Journal of Research in Music*

    Education, 58(2), 168-183. Retrieved from

    http://www.jstor.org.milligan.idm.oclc.org/stable/40666241

Martin J., B. (2007). Performer, rater, occasion, and sequence as sources of variability in music

    performance assessment. *Journal of Research in Music Education, 55*(4), 344.

    doi:10.1177/0022429408317515

McCormick, J., & McPherson, G. (2003). The role of self-efficacy in a musical performance

    examination: An exploratory structural equation analysis. *Psychology of Music*, *31*(1),

    37-50. doi:10.1177/0305735603031001322

McKevitt, C. T. (2016). Engaging students with self-assessment and tutor feedback to improve

    performance and support assessment capacity. *Journal of University Teaching and*

    Learning Practice, *13*(1), 1-20. Retrieved from https://eric-ed-

    gov.milligan.idm.oclc.org/?id=EJ1097250

Oare, S. (2016). Aural image in practice: A multicase analysis of instrumental practice in middle

    school learners. *Update: Applications of Research in Music Education*, *34*(2), 37-44. doi:

    10.1177/8755123314550415

Pellegrino, K., Conway, C. M., & Russell, J. A. (2015). Assessment in performance-based

    secondary music classes. *Music Educators Journal*, *102*(1), 48-55. doi:

    10.1177/0027432115590183

Saunders, T., & Holahan, J. (1997). Criteria-specific rating scales in the evaluation of high

    school instrumental performance. *Journal of Research in Music Education, 45*(2), 259-

    272. Retrieved from http://www.jstor.org.milligan.idm.oclc.org/stable/3345585

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*(3-4),

    207-231. doi:10.1207/s15326985ep2603&4_2

Valle, C., Andrade, H., Palma, M., & Hefferen, J. (2016). Applications of peer assessment and

    self-assessment in music. *Music Educators Journal, 102*(4), 41-49.

    doi:10.1177/0027432116644652

Wesolowski, B. C. (2012). Understanding and developing rubrics for music performance

    assessment. *Music Educators Journal*, *98*(3), 36-42. Retrieved from

    http://www.jstor.org.milligan.idm.oclc.org/stable/41433277

Wheeler, B. A. (2016). *The effect of using a recording of students' performance paired with*

    specific feedback on the self-evaluation calibration accuracy among novice

    performers (Unpublished doctoral dissertation). University of Kansas, Lawrence, KS.

    Retrieved from

    https://kuscholarworks.ku.edu/bitstream/handle/1808/21942/Wheeler_ku_0099D_14759_

    DATA_1.pdf?sequence=1

Yarbrough, C. (1987). The relationship of behavioral self-assessment to the achievement of basic

    conducting skills. *Journal of Research in Music Education*, (3), 183-189. Retrieved from

    http://www.jstor.org/stable/3344960