Assessment in the Choral Classroom: A Case Study of a Secondary Choral Program

Elizabeth (Libby) R. Hearn[1]

Abstract

Assessment in the large choral ensemble classroom continues to be a widely examined topic among music education practitioners and scholars. Scholars have dedicated a significant body of research to identifying and examining the assessment practices of music educators (Denis, 2016; McQuarrie & Sherwin, 2013) However, due to their design, a number of the studies did not thoroughly explore the why and how of music assessment through the voices and experiences of students and teachers (Kotora, 2005; McClung, 1996; McCoy, 1991; Russell & Austin, 2010). This instrumental case study (Stake, 1995) explored perceptions of assessment practices as reported by high school choir students and their choir teacher.

Research questions examined participants' beliefs about assessment, the factors that influenced those beliefs, their experiences with assessment practices, and the challenges of assessing choral music students. Through analysis of multiple types of data collected from various data sources, the findings revealed that the teacher used both musical and nonmusical assessment practices to evaluate student learning. Students perceived all assessment practices, musical and nonmusical, to be in support of what they viewed as the primary goal of the choral program—ensemble achievement. The study identified external and internal influences that directly affected the use of assessment practices at Allen Thomas High School (ATHS), including the choir's role in the school curriculum and culture.

Keywords: music assessment, secondary choral music, grading, evaluation, music education

[1] Department of Music, University of Mississippi, Oxford, MS., USA

Corresponding author:

Elizabeth (Libby) R. Hearn, Department of Music, University of Mississippi, University, MS 38677
Email: ehearn@olemiss.edu

Assessment in music education has been the subject of debate among music education professionals for many years. Paul Lehman's pioneering publication, *Tests and Measurement in Music* (1968), served as a call to music educators who were seeking, at the height of the post-Sputnik era, to understand the importance of measurement and evaluation in music classrooms across the country. But over fifty years later, professional dialogue remains largely unchanged as music educators continue to search for answers to the many questions about music assessment. Through initiatives such as the National Association for Music Education (NAfME) Model Cornerstone Assessments (2016) and revisions of the National Core Arts Standards (2014), music educators have positioned themselves, although with varied success, to argue that music possesses merit as a necessary curricular component in American public schools.

The unique challenges of music teaching and learning have created tensions among administrators, teachers, policymakers, and music educators (McClung, 1996; Russell & Austin, 2010). Pedagogues such as Colwell (2008) and Lehman (2008) posited that music educators generally struggle to find common ground on issues about goal setting and curricular choices. They reported that such controversies have contributed to a variety of views among music educators about the nature of assessment. Such disparity, combined with the subjective nature of music teaching and learning, has impeded attempts to develop fair and reliable assessment measures within the current data-driven, standards-based educational landscape.

Previous studies reported that logistical matters, such as the number of students taught, time constraints, workload, and administrative support influenced music educators when making decisions about assessment (Conway & Jeffers, 2004; Ferm Almqvist et al., 2017; Lehman, 2008; Russell & Austin, 2010). These challenges are confounded by additional challenges, which are unique to ensemble teaching and learning including a lack of professional development on the topic of assessment; discrepancies in contact time with students (due to scheduling issues and/or interruptions); lack of strategies for assessing individual students in large classes; addressing parent and student apathy; and a shortage of available resources for collecting, managing, and storing assessment artifacts and data (Kotora, 2005; McClung, 1996; Russell & Austin, 2010).

With increasing pressure to provide data and to assess individual music students' skills and knowledge, extant literature indicates evidence of prevailing trends toward assessment practices in ensemble music classrooms that do not reflect a student-centered, individualized, data driven approach. LaCognata (2011), Kotora (2005), and McCoy (1988) all reported a prevalence of music ensemble teachers (instrumental and choral) using assessment practices based on nonmusical criteria, especially attendance and participation. According to Russell and Austin (2010), though some music teachers incorporated both achievement and nonachievement criteria into their grading practices, they tended to assign greater weight to the latter. Russell & Austin (2010) concluded that choral teachers assigned greater weight to attitudinal assessments, while their instrumental counterparts prioritized performance assessments.

The inherent subjectivity and fairness issues related to assessing students' behavior, participation, attendance, and attitude, and the widespread use of nonachievement criteria for assessment in music education may leave music educators vulnerable to challenges from students and parents, including legal challenges (Russell, 2011). Moreover, the use of nonmusical and nonachievement criteria-based assessments in music classrooms negatively influenced stakeholders' perceptions of the rigor and expectations of school music programs (Aitchison, 1993; Denis, 2016; Kotora, 2005; McClung, 1996).

These concerns, and the general trends toward greater accountability, have led music teachers to utilize a variety of achievement-based assessment strategies to evaluate musical skill and knowledge including alternative assessments (e.g., portfolios, projects, journals), traditional written exams, and individual and ensemble performance assessment (LaCognata, 2011; Kotora, 2005; McClung, 1996; McCoy, 1991; Russell & Austin, 2010). Even though such strategies were designed to objectively measure and evaluate individual musical achievement, written exams may be vulnerable to questions of validity (Wesolowski, 2020) and bias (McMillan, 2018) while performance assessments are reportedly unreliable (Bergee, 2003; Latimer et al., 2010; Reimer, 2009; Ryan & Costa-Giomi, 2004). Additionally, performance assessments of group achievement have not demonstrated reliable evidence of individual achievement (Broomhead 2001; Henry & Demorest, 1994).

These studies suggest that a significant body of research has addressed the many issues related to music assessment. However, most of the studies, due in part to their designs, did not explore the why and how of music assessment through the voices and experiences of students and teachers. Though several studies have examined teachers' use and perceptions of assessment in music education (Harrison et al., 2013; Hawkins, 2018; Kancianic, 2006; McCoy, 1991; McClung, 1996; Reimer, 2009; Russell & Austin, 2010; Tracy, 2002), fewer studies have investigated student perceptions of those assessment practices (Aitchison, 1993; Conway & Jeffers, 2004; Kotora, 2005; McClung, 1996). Furthermore, the existing research fails to address the thought processes of teachers when planning and implementing assessment practices in their classrooms.

This qualitative perspective could provide additional insight to previous findings. The lived experiences of this study's participants combined with the findings generated from the field may contribute to a growing body of knowledge that seeks to understand the values, beliefs, and perceptions of those engaging with and making decisions about assessment in music classrooms. Therefore, the purpose of this instrumental case study was to examine participants' experiences with assessment practices in a secondary (high school) choral ensemble class and to explore their values and beliefs about assessment.

To do so, I used a case study methodology to collect and analyze data to address three primary research questions:

1. How do participants perceive, value, and experience assessment practices in the choral classroom?

2. How do school culture and other factors influence participants' perceptions, values, and experiences of assessment in the choral classroom?

3. What challenges do choral teachers and students encounter when implementing and engaging with assessment in the choral music classroom, and how do those challenges influence assessment practices?

Method

I used an instrumental case study (Stake, 1995) to investigate assessment practices as they existed in a real-life context. The purpose was to investigate a key topic or concern about a single, bounded case leading to understandings and assertions about various assessment practices in the choral classroom. By situating myself in the naturalistic and interactive learning environment, the case-study methodology allowed me to create a real-life portrayal of the case and to use field data to examine more thoroughly the key instrumental focus of the study.

Research Site

The site was Allen Thomas High School (ATHS - pseudonym); a suburban, Southeastern public high school in the United States. A total of 212 students were enrolled in choral classes at ATHS at the time of data collection. The choral curriculum included five developmentally sequenced choral music classes: Mixed Choir (beginners and advanced mixed, grades 10–12), Men's Choir (grades 9–12), Women's Choir (grades 9–12), Concert Choir (women's choir, grades 10–12), and AT Singers (advanced mixed, grades 10–12). All ATHS choirs were curricular. Auditions, which took place in the spring prior to the next academic year, determined students' choir placements.

Many of the choir students experienced formal music instruction for the first time as a student at ATHS. The feeder elementary and middle schools that ATHS students attended did not offer music instruction by a certified music educator. Table 1 on page 45 provides demographic information for ATHS as reported by National Center for Education Statistics (2018) compared to the enrollment in the choral classes. I used student registration information provided by ATHS. The research site provided variation in the demographic makeup of each choir including the socio-economic differentiation of the student population, and a representation of students at various stages of their musical and vocal development.

Table 1

Demographic Information for Andrew Thomas High School and the Choral Program

	Total students	American Indian	Asian	Black	Hispanic	White	Two or more races	Male	Female	% Free and Reduced Lunch
ATHS	1520	>1%	>1%	35%	7%	56%	>1%	49%	51%	36%
Men's	33	0%	0%	9%	>1%	90%	0%	100%	0%	N/A
Mixed	54	0%	0%	20%	0%	79%	0%	41%	59%	N/A
Concert	38	0%	0%	15%	>1%	84%	0%	0%	100%	N/A
Women's	58	0%	3%	17%	3%	74%	>1%	0%	100%	N/A
ACSingers	17	0%	0%	23%	0%	76%	0%	41%	59%	N/A

Participants

Teacher

Ms. Andrews (pseudonym) was the lead choral educator at the high school and had been teaching at the school for 13 years. She developed a large choral program built on what she described as a "foundation of musical skill and understanding." According to the ATHS choral handbook, Ms. Andrews desired to "cultivate the development of skills in vocal technique, sight-reading, vocabulary enrichment, and performance skills." In her role as lead choral teacher, she oversaw and administered all forms of classroom assessment procedures.

Students

Students (see Table 2 on page 46) were selected to participate based on their availability and willingness to contribute to the study. I aimed to gather a representative sample of the diverse student population. Student participation represented a cross section of the school and choir population and provided perspectives from varying levels of experience in choral singing, vocal skills, and musical development. A total of 20 students participated: 10 students were beginners and 10 were advanced; 10 students identified as male and 10 students identified as female. The advanced students, mostly 11th and 12th graders, possessed more experience (three to four years) and greater familiarity with the assessment practices. For beginning students (mostly 9th and 10th graders), this was their first or second year in a choir.

All 20 students participated in one of three focus group sessions. Following the focus group discussions, I selected the 10 individual interview participants. I selected students based on my judgment of their multifaceted contributions to the focus group discussion and their ability to inform the research questions.

Table 2

Student Participants

Name (pseudonym)	Grade	Choir	Level	Years of Experience	Gender	Race
Brandi	9	Women	Beginner	0.5	Female	White
Alice	9	Women	Beginner	0.5	Female	Black
Miranda*	10	Women	Beginner	1.5	Female	White
Maria*	10	Concert	Beginner	1.5	Female	Black
Nina	10	Concert	Beginner	1.5	Female	White
Jake*	9	Men	Beginner	0.5	Male	White
Sidney	9	Men	Beginner	0.5	Male	Black
Lane	10	Mixed	Beginner	1.5	Male	White
Charlie*	10	Mixed	Beginner	1.5	Male	White
Jude	10	Men	Beginner	1.5	Male	Black
Jess	12	ATS & Mixed	Advanced	3.5	Female	Black
Catherine*	12	ATS & Mixed	Advanced	3.5	Female	Black
Liz*	11	ATS & Mixed	Advanced	2.5	Female	White
Barbara	12	ATSingers	Advanced	3.5	Female	White
Becca	11	ATS & Mixed	Advanced	1.5	Female	White
Tommy	12	ATSingers	Advanced	1.5	Male	White
Mark*	12	ATS & Mixed	Advanced	3.5	Male	White
Joseph	12	ATS & Mixed	Advanced	3.5	Male	White
Michael*	12	ATS & Mixed	Advanced	3.5	Male	White
Bryant	11	ATS & Mixed	Advanced	1.5	Male	White

Note. * denotes participants who completed an individual interview

Data Generation

Prior to collecting data, I received Institutional Review Board (IRB) approval and collected informed consent and assent forms from all participants. Data generation occurred over a 10-week period from October 2018 to March 2019. Data were generated from multiple sources. My role was participant observer where participants understood my function as researcher (Stake, 1995).

Primary data sources included field notes generated during observations of assessment practices, focus group discussions, teacher reflections, and semi-structured interviews (Roulston, 2010). Focus group and individual interview protocols are available in the Supplemental Materials. I included and analyzed a collection of artifacts, such as handbooks/syllabi, grade reports, and written student work.

Teacher data included five written reflections and two semi-structured interviews (Roulston, 2010). The written reflections were responses to prompts, which I emailed to Ms. Andrews and were generated throughout data collection. The purpose of gathering teacher data was to better understand her choices, values, and lived experiences: the 'how and why' dimension of implementing various assessment practices. Data generated from the students included the content of focus group discussions, individual interviews, and artifacts of student work.

Data Analysis

Analysis of the data included the interpretation and deconstruction of not only the data in research texts, but also the meaning of my impressions (Stake, 1995). I transcribed observational field notes and teacher reflections into texts throughout the data collection period. I recorded and transcribed all focus group discussions and interviews. I completed transcripts before subsequent interviews. Groups were scheduled to use the data to inform the direction of future protocols.

Both during and after data collection, I employed categorical aggregation, using NVivo, to deconstruct and organize complex data into categories as I searched for meanings, patterns, and relationships that supported behaviors, issues, and the contexts unique to this particular case (Stake, 1995). At the conclusion of my fieldwork, I used Maxwell's (2013) strategies for qualitative data analysis as a guide: (a) reading and memoing, (b) categorizing strategies (coding and thematic analysis), and (c) connecting strategies (narrative analysis) (p. 105).

Trustworthiness

I pursued credibility of the data and interpretations through triangulation, the gathering of rich data, prolonged engagement, and member checks (Stake, 1995). Triangulation of multiple types of data (audio recordings, memos, jottings) collected from different sources (artifacts, interviews, and observation) sought to verify the findings and interpretation. My

prolonged engagement with the participants and site contributed to a greater understanding of the various complexities of this case, thereby strengthening the trustworthiness of the findings (Stake, 1995). Through the process of member checking, participants examined rough drafts of interview transcriptions and reviewed them for accuracy and agreeability (Stake, 1995). Ms. Andrews also approved the findings and the final manuscript.

Clarifying Researcher Subjectivity

My relationship to the site of this case study was multilayered. The teacher was a colleague and friend. I also worked at ATHS as a collaborative pianist. While there were benefits to my familiarity with this setting, such as trust and rapport with the students and teacher, that familiarity held the potential to create power or boundary issues that could impact my study. The in-depth nature of the study could potentially create tension if participants viewed findings as a reflection of my personal feelings towards them as individuals.

To address the potential power issues in my new role as a researcher, as opposed to my typical role of teacher/instructor, I was deliberate and intentional about how I defined my role. The recruiting script, the interview script, and my interactions with the participants communicated my desire to create a comfortable, confidential, and safe space. Individual student interviews and focus group discussions took place away from the view of Ms. Andrews and other peers. Students could skip interview questions that they were not comfortable answering. I also gave students the opportunity to review and approve/disapprove of their contributions. To address the existing adult-student power differentials, I maintained respect for the participants, I listened attentively, and I responded to their statements from a neutral and empathetic position (Roulston, 2010).

The student participants and Ms. Andrews appeared comfortable and eager to talk with me. Ms. Andrews reassured the student participants that she supported the study and encouraged them to be open and transparent without fear of consequences. Therefore, the familiarity between myself and this case ultimately proved to benefit the depth and authenticity of my findings.

Findings

Descriptions and narrative illuminated aspects of the students' and the teacher's lived experiences with the various assessment practices. Themes emerging from data analysis included the following: ensemble achievement, individualized assessments reinforcing accountability, the choir's role in the school curriculum, and challenges that impacted assessment. External and internal influences at ATHS, including the perceptions of the learning outcomes and nature of the choral music experience, interacted to influence the assessment decisions and practices. It is important to situate these findings in the context of the formal assessment practices implemented at ATHS to draw conclusions from the emergent themes .

Table 3
ATHS Assessment Practices

Assessment	Description	Graded (Yes/No)
Individual Participation	Students received a daily participation grade. Students received 50 points per week for participation and points were deducted from that total score when students failed to meet the expectations communicated in the syllabus. Expectations included: (a) actively participate in class each day, (b) remain quiet when not singing, and (c) have only required materials at seat. Ms. Andrews posted summative, weekly participation grades and reported this information to students and parents through an online data collection and reporting platform used at ATHS to track attendance, record, and report grades, and communicate with parents.	Yes
Concert Attendance	Students received deductions from their concert attendance grade for unexcused absences from concerts, tardiness to the performance, incorrect uniform, and improper concert etiquette (behavior). Students who were excused from a performance were expected to report to the teacher to schedule make-up work, which included written work or an aural exam. For each required performance, Ms. Andrews assigned parent volunteers to document student attendance and to evaluate their concert attire. Ms. Andrews evaluated the documentation to determine the concert grade and was available to the students upon request.	Yes
Contextual Singing Assessment (CASA)	Students used a personal device to generate an audio recording of themselves while singing and performing with their classmates as an ensemble during class rehearsal. Students submitted recordings via Charms or email. Each student received individualized, formative feedback from Ms. Andrews. Students were occasionally required to submit a self-assessment of their CASA recording.	No
It Works in Theory (IWIT)	IWIT (Heron, 2013), a standards-based assessment model, utilized differentiated and peer instruction to teach sequential and comprehensive music theory concepts to students in ensemble classrooms. Students moved at their own pace through 26 levels of theory knowledge. To advance to the next level, students were required to score a 90% on the theory exam. The student's nine-week theory grade was an average of all theory test scores.	Yes
Other Singing Assessments	Students were individually assessed using live solo and small group performances during rehearsals. Examples of singing assessments observed included: (a) performing in quartets, (b) individual demonstrations in class, (c) performing in octets, and (d) recording small groups and playing back for group evaluation. Students in AT Singers and Mixed Choir were occasionally required to submit recorded assessments that were completed at home or in a practice room (in contrast to CASA).	No

Assessment as Essential for Ensemble Achievement

The students' and their teacher's definition of what it meant to sing in a school choir held a strong association to the purpose of assessment. The concepts of team, collaboration, and a focus on interactive learning were foundational to the choral curriculum and learning environment at ATHS. The students' views about assessment reflected those values. Although the teacher had individual learning goals for students, the goal of ensemble achievement (performance goals accomplished by the collective efforts of the ensemble) reflected the spirit of what it means to sing in a choral ensemble and, therefore, emerged as the primary goal. A central finding of this study was that students understood individual achievement and assessment as being linked to ensemble achievement and in some ways secondary to ensemble achievement.

> When it all comes together, it just is awesome. If you don't grow as an individual with your understanding of music and your knowledge of the particular piece, then you're not going to contribute anything to the ensemble, and then the ensemble is not gonna [sic] grow as much. So, it's kind of like building blocks, one has to have them before the other one. Both [individual achievement and ensemble achievement] are yes, very important. The end goal is going to be ensemble growth, but the short-term goal is gonna [sic] be personal growth. (Liz)

Liz understood her individual achievement in choir as directly related to ensemble achievement with assessment being a component of the process, not the goal. Students, like Bryant, aspired to excel as individuals so that they could be valuable members of the team (ensemble):

> And at the same time, with other classes' homework and stuff, you do it really reluctantly. You don't really want to do it, but with choir, the work that we take home, practice and stuff, you want to do it because you want to make it better. You want to do it for the overall outcome, not just yourself. (Bryant)

Being a part of something bigger than themselves was a key component of their learning experience that influenced their perceptions of assessment and their individual achievement.

Role of Participation

Ensemble achievement, as the primary goal, required students to develop a fundamental understanding of and to demonstrate mastery of the proper rehearsal and performance behaviors and skills. These behaviors and skills were perceived as necessary to cultivate an environment that was conducive to the type of interactive learning and performance that is typical in choral music education.

Ms. Andrews regarded the participation assessment as useful, but she acknowledged the issues that were potentially associated with this practice by stating: "There's no way I can monitor each and every child's participation fairly." She attempted to be transparent and open about the participation grades and to address subjectivity by posting a daily participation log. Furthermore, she was candid about the difficulty in accurately assessing students' individual daily participation: "Who's to know if they just start singing the minute that I walk by and then they stop … you can't really know their level of participation from the conductor's podium."

Her response reinforced the complexities of evaluating participation in a large group setting. Like many other music educators, Ms. Andrews used a combination of "subjective impressions and objective documentation" to assess daily participation (Russell & Austin, 2011, p. 44). Because student engagement and attendance were essential to the achievement of the ensemble, participation and attendance were heavily weighted (68% of the students' nine-weeks grade).

When I asked students how they were assessed in choir, all three groups responded with participation first. Students, such as beginning student Maria, understood the expectations for participation:

> The 50 [grade] in choir means you come to class; you participate. You're not just sitting in your chair, talking to your friends, disturbing them from learning the music that they need to learn because you also need to learn it, too. You had to participate in class, be on time, you had to show up. You had to act like you're into it, instead of just sitting there like: "Oh, this is boring and I'm ready to go home." And I sang when we were supposed to, and I put effort into the class. I didn't just sit there and slouch back in my chair, barely open my mouth when I sang.

Students like Maria understood there were observable expectations for meaningful participation, including but not limited to punctuality, singing, proper posture, responsiveness to the director, and attentiveness. Catherine, an advanced choral student, described her participation grade as including musical outcomes, including addressing or fixing musical mistakes:

> That means that I put forth effort to do whatever we're doing in that week. I put forth effort to achieve. When she says, "Fix something," fix it, not keep messing up that same thing that she said fix and just doing what she says to do when she says to.

Other students shared their perceptions of the expectations for their participation evaluation:

> I'm thinking, just being on your phone, maybe a lot [*sic*]. I would think that that would come off. Or just if you're obviously not reading music or singing or not paying attention at all. I think as a teacher, you can somewhat tell. You can't necessarily tell if someone's into it, but you can tell if someone's not into it because their body language will show you. (Michael)

For the students, the participation assessment practice was a necessary component of the choral experience. It reinforced the classroom structure and dynamic that was essential for cultivating ensemble identity and achievement. Therefore, student participants did not perceive the evaluative process that resulted in the participation grade to be unfair.

Evaluating Attendance

Ms. Andrews formally assessed concert attendance to motivate students to attend performances. She stated:

> Concert. That one is just, you got to come to the concert. It's [concert grade] to get them to come there because if they come there, they're going to do all the things or most likely, they're going to do all the things that they're supposed to do at the concert.

By assigning a grade for concert attendance, she reinforced the practical and tangible aspects of presenting a musical performance that require each ensemble member to be present. Therefore, the grade reflected the significant value of the performance as part of the choral class experience. Students understood the expectations for concert attendance and concert attire:

> The Christmas concert I showed up on time and I was wearing my choir dress. My hair was down and out of my face. I had pretty makeup on and didn't look like I rolled out of bed. I had my earrings, my pearl necklace, and my closed-toe black shoes, and I was ready to sing. I was there on time. (Liz)

Maria stated:

> It says that you are willing to put yourself forward and come and be in the choir, instead of just sitting at home, when you know you're in choir, you have a concert. So, you have to push yourself, be like, "Oh, wait. I have a concert tonight." You just can't stay at home. You have to go to your concert, 'cause [*sic*] that is a huge grade.

Mark described what he perceived to be Ms. Andrews's evaluation of the students' individual contribution at concerts: "Well, whenever you walk in you sign in and you should be wearing your entire uniform. So that's at least 100 right there." Interestingly, students' musical contributions on stage were not evaluated as a part of the concert assessment, but the value associated with concerts shaped Mark's understanding of the criteria for this concert assessment. He stated:

> …Then to the best of your ability, if you're on stage, if you're engaged watching her conduct. 'Cause [*sic*] if you're just standing there doing nothing she's gonna [*sic*] see you. We're choir kids, we pop around and jump. We move with the music. And if you're doing nothing, she can see that. (Mark)

Concerts and performances were a natural outcome of the choral music experience. For that reason, students like Lane were motivated intrinsically (being a valuable ensemble member) and extrinsically (grade) to attend:

> We work for months and months, so the performance, the concert, that's like the most important thing to go to. That, that's what should be graded the highest. We work the hardest and the longest on [concert music]. We spend the most time on it. (Lane)

For the students, concert attendance held similar value to that of a final exam. Concerts represented the culmination of their daily work and achievement as individuals and as an ensemble.

Individualized Achievement Assessments and Student Accountability

Ms. Andrews utilized several individualized, achievement-based (musical) assessment practices focusing on theoretical musical knowledge and singing skill. Though these individualized assessment practices measured and reported individual student achievement, students recognized that they also influenced ensemble achievement by holding individuals accountable for their own contributions.

Accountability Through Music Theory Assessments

The self-paced, sequential, and comprehensive attributes of the It Works in Theory (IWIT) model (Heron, 2013) distinguished it as an innovative approach to assessing music theory knowledge. Ms. Andrews believed that students saw the theory assessments as more academic because these written, formal assessments, generated an individual grade, a grade that motivated many of the students to continue to progress through the theory program.

While the exam format was a traditional pen and paper test, the timing and frequency of the assessments were not traditional. Ms. Andrews desired to find additional ways to

integrate theory concepts into the rehearsal process. Her statement here reflects tensions experienced by other music educators when balancing instruction, assessment, and transfer of skills:

> I don't want to compartmentalize it, but I do sometimes. I was just thinking yesterday, I have to do a better job of referencing these things that they're learning when we're talking about theory inside of our rehearsal and making sure that they understand that those things cross over into the stuff that we're doing. That's hard to do just because we're not doing those things together at the same time. We have theory time and then we rehearse, which I know is not the best, but how else? I don't know how else to do it at this point.

Students saw the theory grade as consequential, affecting their achievement in choir, and as a qualitatively different component of their choral music class, more akin to assessment in other high school classes:

> Well, the motivation, with theory, for me considering I'm not a music major, I'm not gonna [*sic*] be in music education, the motivation is less like, "Ooh, this is exciting," and more like, "All right, for the grade and for being able to pass and do well in the class." That's kind of bad, but that's how it is with other classes. That's how I am with a math test. I understand why it's necessary, but it's not necessarily the fun part of it. (Michael)

Despite Ms. Andrew's concerns that she compartmentalized theory in her curriculum, students were able to articulate specific knowledge and understanding gained through IWIT that transferred to their ensemble experience:

> Now that I know some about theory I can be like, oh this is this key signature, which is kind of cool. I just think that's cool. That's just me being a choir nerd. Then other times I can be like, oh this marking right here means attack it and release, and then this dynamic marking means forzando, so loud, quiet, loud. So, it's cool. I can put more into the music than other people might be able to because I have a better understanding of what is written into [sic] the paper. (Liz)

Additionally, IWIT generated concrete student achievement data that Ms. Andrews used to report student learning to administrators and parents. This formal, objective assessment practice distinguished IWIT from the other strategies used to assess the students.

Accountability Through Singing Assessments

Performance assessments included various types of singing assessments. These assessments were almost entirely focused on evaluating individual performances of the repertoire

that was being rehearsed in preparation for performance, as opposed to solo literature or sight-singing exercises. Ms. Andrews implemented several forms of live solo and small group (one singer per part) singing assessments where students demonstrated proficiency on their part.

The use of a contextual, authentic singing assessment (CASA) was different from the more conventional forms of performance assessments used in that students recorded themselves singing within the context of the ensemble, not away from the ensemble as demonstrated in the process for collecting the solo, recorded singing assessments. That the students recorded in the choral context and could hear other voices while recording offered unique benefits and removed some of the anxiety experienced when singing alone in front of their peers. These recordings offered additional benefits for rehearsal planning as Ms. Andrews explained:

> I have started to use [CASA] when I hear something individually that I don't hear in the group, but it's from a lot of people. Then I can go back and say, "Okay. The basses don't really know that measure. It sounds like they do because I'm hearing two or three sing the right notes, but the rest of them are off." I can go back and fix. That's what happened with the last one that they did. They didn't know [the measure]. I thought they knew it, but they didn't know it.

Ms. Andrews believed CASA to be the most valuable assessment practice she used; it generated data that demonstrated and tracked musical growth and individualized achievement while also reflecting the daily learning environment and context. The students agreed:

> I feel like you can definitely hear me even with the other parts in the background, and that probably personally helped me, the other parts. 'Cause [*sic*] that's just how I'm used to singing something, but that still didn't necessarily come out perfectly. (Michael)

Liz reflected on her experience with CASA:

> I feel like I sound a little bit better than that [recording]. It feels different because when you're singing with a group you can hear how your part blends in with all the other parts. And I feel like it's easier when I'm singing with the choir, because I don't have the melody there because the guys do, so I can hear how my part supports their part.

The data generated by CASA facilitated a more accurate assessment of the students' individual contributions. Ms. Andrews stated, "That's what they really are doing. I mean, they're not singing a solo." She also perceived that CASA had an immediate effect on en-

semble achievement:

> I noticed that the ensemble sounded better. I think everyone was attempting to make their best sound, which they do not always do. I also heard some voices singing louder than normal. I think it's louder because they're all singing, trying to get it into the [device], so it records, but I think some things are better actually because they're thinking about them. They know that they have to, okay, I just said, "I'm listening for right notes, right rhythms, good tone, and good diction." Okay. They're thinking about those four things and trying to do them correctly. (Ms. Andrews)

Several students were aware that they sang differently when recording their solo assessments compared to the CASA recordings. Catherine stated: "Because when I'm by myself, sometimes I'll sing it differently, a little bit, then you would in a choir, because in a choir, you're all supposed to fit the same sound that you're looking for." Another student commented that she could sing "freely" when recording a solo singing assessment.

Students also perceived that recording themselves within the context of the ensemble was helpful for them and Ms. Andrews, less stressful than the solo recording, and that it improved their individual understanding of the repertoire:

> I think it helps, because you're singing into the phone and then sending it to Ms. Andrews so she can individually assess you instead of doing it as a group where she hears everybody and not just one person. So, then she knows, hey, you need to come up on pitch, you're too high so you need to drop down, or if you're right on. (Jake)

Ms. Andrews implemented individualized musical assessments to foster the individual development of students, which contributed to ensemble achievement. Although students rarely performed as individuals in class, individualized singing assessments were important to the students and to Ms. Andrews. The use of performance assessments emphasized the importance of individual musical skills and reinforced the value of everyone's contribution.

Choir's Role in the School Culture and Assessment

"Choir is different" was a sentiment communicated frequently by students when describing choir class and how it compared to their other classes. Students integrated this distinction into most facets of their choral music experience. The participants' beliefs and values about the nature of the choral music experience, influenced by the choir's role in the school curriculum, were instrumental in understanding how assessment fundamentally worked in this setting.

Choir was an elective course at ATHS. Participants felt that this designation influenced

the perceived role of choir in the school culture. Ms. Andrews was acutely aware of the elective designation when making decisions about outcomes associated with learning in choir and the assessment of those outcomes. Students were required to take one fine arts credit (elective) for graduation and they described choir as somewhat extracurricular, even though it occurred during the school day. Liz explained: "You're expected to learn that [math], but you're not expected to take choir unless you want to, or you need a fine arts credit." Bryant noted: "Choir is extracurricular. And [parents are] like, 'That's what it's supposed to be extra. It's not the main thing.'"

Students cited various reasons for electing to participate in choir: they enjoyed singing, they liked the social benefits of being in choir, choir was "fun," it was less "stressful" than other parts of their day, and they liked the teacher. These reasons contrasted with their descriptions of their required courses and reinforced the differences between participation in choir and other classes in which one student described learning as "being hunkered over a desk." Because the learning environment in choir was more interactive and group-oriented than their required courses and even other electives, participants believed that the assessment practices should reflect those differences.

Choir as Activity

Much of the learning in choir occurred within a group setting, and students demonstrated learning through the psychomotor process of singing. Therefore, some participants believed choir to be more like an activity, such as athletics, and less academic in nature. Ms. Andrews stated:

> Everyone thinks that choir is not a real class. . . most teachers think "oh, you guys are just singing." Like, "All y'all [*sic*] are doing is singing." I would think that most parents think that. I think the general public thinks that we don't have a curriculum really and we just sing.

These perceptions troubled Ms. Andrews. She structured her classes to provide sequential learning opportunities appropriate for students with varying levels of musical skill and knowledge. And yet, to her, her colleagues, and the public perceived singing to be of less academic value or rigor than other types of acquired skills and knowledge.

The belief that choir was an activity and not an academic class was evident in how students described the ways they were assessed in required courses compared to how they were assessed in choir. Students shared that formal, individual assessments occurred less frequently in choir, when compared to their other classes, contributing to a less stressful learning environment:

> I feel like people don't have as much stress in choir as other classes, because with your other classes you have to study for all these big old tests, and you have to be

doing your homework at home. If you're not getting something, you have to figure it out. But in choir, there's not that stress of, "Oh, well I have this test on this day," every week. (Maria)

Students like Michael saw their experience in choir as having similarities to other activities, such as sports, and less like their required, academic classes, such as math and history:

You actually are going and talking about [choir] almost, how my brother would [about] a football game. It's like: "All right, we're learning this and this and I'm excited to sing this at this time." And then Josh will be like: "Yeah, and we did this in practice," but it's not like the grade of it all, even if there is a grade. (Michael)

One student, Mark, was surprised when I asked him if participation was assessed in his required courses: "Like a participation grade in other classes? It's just strictly academic so you don't get judged on whether you're paying attention or not or else you'd get zeros all the time. [Choir] is an elective." Mark and others considered participation to be an irrelevant form of assessment in required classes because grades were determined by individual achievement on formal, mostly written, assignments. Individual participation in core academic courses did not impact the collective achievement of the required class as it did in choir.

Principals, Parents, and the Public

Administrators' (local and district level) expectations for assessment in choir and their position that choir was not academic influenced Ms. Andrews's beliefs about assessment. During our interview, she reported the lack of interest and expectations from administrators when it came to assessment in choir:

Ms. Andrews: The fact that from the top down, everyone wants us to keep records of everything and we need data. Everybody is talking about that… they are not talking to us because nobody expects us to do anything.

Researcher: Have they ever said that? Or is it just understood?

Ms. Andrews: I think it's just kind of understood.

Researcher: And there's never been a time where they've asked for [data]?

Ms. Andrews: No, there's never been a time where anybody has asked for it. But there actually has, you know, been times when I've gone to my principals and said: "This is what I'm doing, this is what you're asking

from me, but I don't really have a way to give it to you." [They said:] "Oh, don't worry about it. It's ok. We just really want that from English, math, science, and social studies."

Teachers of required courses and other content areas were expected to generate individual data to document student learning and were given financial support to do so. The lack of expectations, investment, and guidance for individualized assessment and data reporting for choir influenced Ms. Andrews's beliefs and ultimately her decisions about assessment in choir.

According to some students, parents differentiated between required classes and choir class in their expectations for levels of achievement, which influenced assessment values. Most students indicated that their parents prioritized their performance in academic classes as more important than choir. For some students, participation in choir was a privilege given to those who were doing well in their required courses.

Students reported that their parents made participation in choir contingent upon meeting their expectations in other courses. For example, Bryant shared: "My parents have told me straight to my face: 'You're doing academic homework first because choir's less important than the academics.'" Liz explained that she was required to get good grades to be able to participate in extra choir activities: "My parents have said to me that if I don't get good grades in my academic courses then I can't go on to All-state or stuff like that."

The pressure to present quality performances was deeply engrained in the music education culture at ATHS and in the community. These public performance expectations influenced the scope and purposes of assessment. While Ms. Andrews believed that assessment was necessary, the performance demands she experienced limited the frequency of formalized assessments. The visibility of the choir, through well-attended public performances, reinforced the value of participation by regularly putting them on display. High-quality choral performances, a demonstration of ensemble achievement, shaped the public's perceptions of the choral music program.

The elective designation and the role of choir in the school culture and curriculum was a powerful distinction that influenced the students' and Ms. Andrews's expectations of assessment in the choral classroom. Ms. Andrews supported the notion that while external and internal factors affected her assessment decisions, in turn, assessment (or lack thereof) influenced the way others perceive choir:

> I was going to say, choir compared to other classes, I think in just over my time as a teacher, I have seen that a lot of other teachers, administrators, people in education don't view choir as a class, maybe, because of our lack of true, individualized assessment.

These beliefs, along with choir's non-academic status, voluntary participation, and focus on ensemble identity coalesced to create a unique learning environment. Participants per-

ceived the assessment model to be effective and appropriate in this environment.

Balancing Assessment Challenges and Performance Expectations

Instructional time, large enrollment, performance expectations, assessment knowledge and training, and the efficiency of the current assessment practices were key challenges that affected assessment. Ms. Andrews attempted to reconcile the need for individualized assessment practices with the demands and challenges that existed in her classroom:

> I think that I'm not satisfied with the way that I'm currently assessing because I'm not able in the class time to do what I want to do and teach them all of the things that I have to teach them to catch them up from not having music from kindergarten to 8th grade. . .I can't do [individual assessment] every day because we don't have time. If we had every kid sing every song by themselves every day, we could not get through the literature that we're working on.

Like other music educators, Ms. Andrews expressed frustration with the lack of training and experience with individualized assessment practices appropriate for performance-based classes. When asked about how much training she received in her pre-service and in-service professional development she stated:

> I would say not much. Maybe, a couple of days in our undergrad we spent talking about that, I think, but, not related to how to do it in a choral setting. More like the data assessment part of it. And then, we didn't talk about it [assessment] at all in my master's degree that I can remember. And professional development, one session, I think.

She acknowledged that this deficit in her training influenced her early assessment decisions and that she adopted assessment practices that she had learned from student teaching that were solely based on attendance. She has since tried to adopt assessments that were "more reflective of what the students are actually doing in class."

Ms. Andrews reflected that the amount of time required to evaluate singing assessments was a considerable challenge. Due to large enrollment, the number of student submissions was overwhelming:

> It is difficult to try to listen to recordings for each of the 212 students in the choral department all at once. It takes hours and hours to do. I may need to rotate the recorded assessments, maybe do one class per week, so that I don't have to spend so much time grading all at once.

Both the students and Ms. Andrews acknowledged that time combined with large enroll-ment was a deterrent for individually assessing students because more students equate to more time spent evaluating submissions and potentially more time away from the rehearsal process. Though Ms. Andrews desired to individually assess students, the amount of time required to do so with some consistency conflicted with the primary goal of using instruc-tional time to rehearse and prepare music for performances.

Discussion

The perspectives of the student participants and Ms. Andrews illuminated the tensions, outcomes, and challenges of assessment in choir (see Figure 1). Collectively and inde-pendently, three categories of influences impacted assessment at ATHS: (a) external in-fluences, (b) the choir's role in the school curriculum, and (c) internal influences. Based on those influences, Ms. Andrews designed assessment practices which evaluated individual participation and individual musical skills and knowledge.

The ATHS assessment model reinforces the relationship between assessment and ensem-ble achievement in that the assessment practices supported and were essential for accom-plishing the primary choral program goal of ensemble achievement. Ensemble achieve-
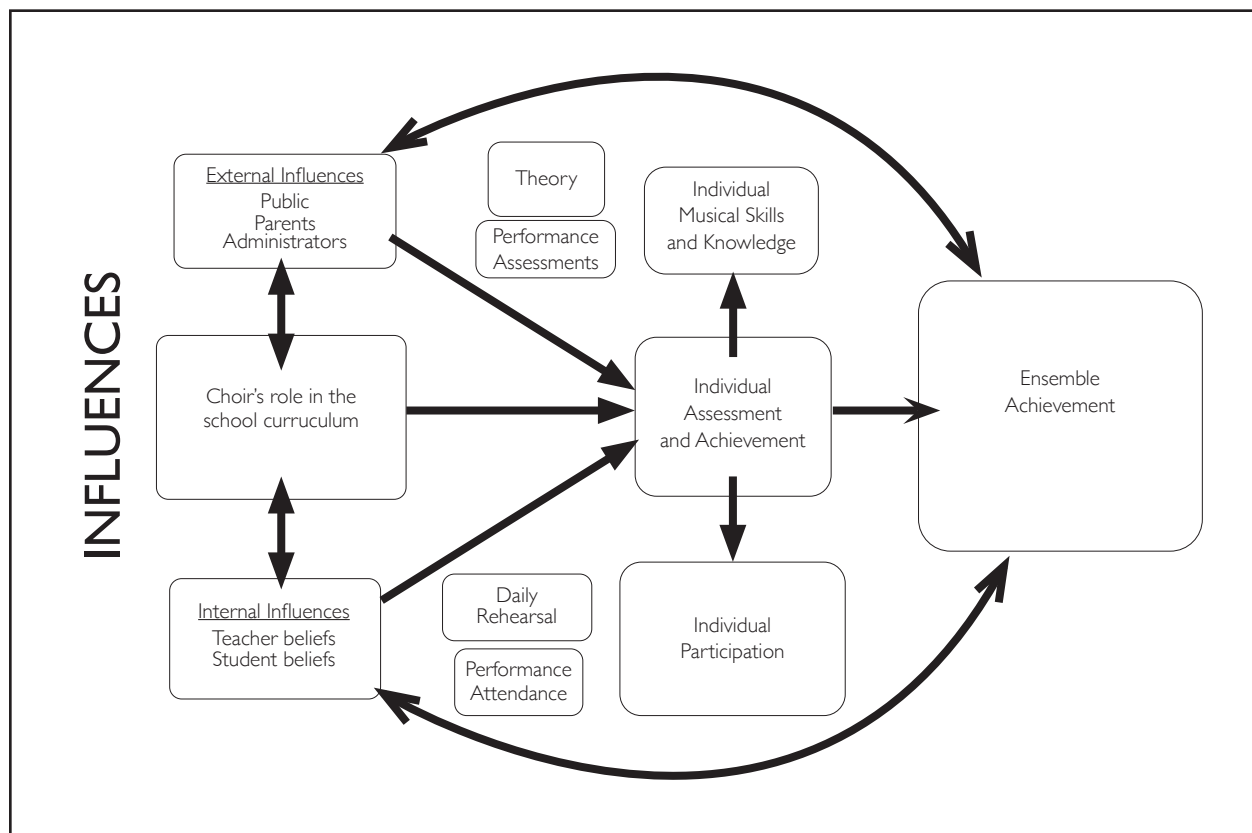

Figure 1

Interactive Model of Assessment at ATHS

ment, demonstrated through public performances and the experiences of the participants, was then instrumental in shaping the role of choir in the school curriculum and the external and internal perceptions of the choral experience. The interaction of these factors, influences, and outcomes forms the synergistic relationship.

Because ensemble achievement and group instruction were highly valued, the use of nonmusical assessments for individuals was considered appropriate by participants and was also perceived to be a necessary component of the complete choral music experience. This finding contrasted with McClung (1996) who reported that many choral students preferred music tests and sight-reading assessments over attendance and participation.

Though scholars have noted the inherent reliability, validity, and fairness concerns when using nonmusical criteria to evaluate students in large group settings and the potential negative implications for music educators (Harrison et al., 2013; Russell, 2011), students in this study did not perceive there to be bias or fairness concerns with these nonmusical criteria contributing substantially to their final grade. This finding also differs from the results from Harrison et al., (2013) who found that nonmusical assessments led to perceptions of favoritism and issues of fairness.

Russell and Austin (2010) labeled assessment criteria designed to evaluate participation, attendance, and behavior as "nonachievement criteria" (p. 39). However, the findings of the present study cast doubt on labeling these assessment strategies as nonachievement criteria. In this case, students were demonstrating observable skills and understandings about the rehearsal and performance process, namely, voice building, posture, facial expression, engagement with the musical score, and other behaviors that were related to their individual contributions. These skills and behaviors demonstrated achievement and mastery of the expectations for meaningful participation in a choral ensemble.

Despite reported reliability concerns about the use of performance assessments to document students' musical achievement (Bergee, 2003; Latimer et al., 2010; Reimer, 2009; Ryan & Costa-Giomi, 2004), the individual accountability generated by singing assessments at ATHS provided a positive impact on the various ensemble's level of achievement. Like Crochet and Green's findings (2012), participants perceived that the CASA strategy enhanced individual musical skills for performance. CASA also provided information about students' achievement in relation to the authenticity of the teaching and learning process. The validity and reliability of the contextual singing assessment (CASA) warrants further exploration.

That choir was different from required courses and even from other electives was an important finding impacting the use of assessment and perceptions of assessment at ATHS. These findings aligned with previous research which suggested that this distinction influenced administrator and student expectations (McClung, 1996), influenced parent's perceived expectations for grading (Conway & Jeffers, 2004), and shaped the teacher's philosophy of assessment (Tracy, 2002).

Finally, a significant assessment challenge for Ms. Andrews was lack of instructional time, which is consistent with previous findings (Conway & Jeffers, 2004; Kotora, 2005; Russell

& Austin, 2010; Tracy, 2002;). The students' lack of music instruction prior to their participation in choir at ATHS compounded this issue. Ms. Andrews felt pressured to use instructional time to teach and reinforce basic musical concepts and skills while also preparing for performances.

Limitations

While the findings of the current study are valuable to practicing choral music educators, they are limited to a particular group of students who auditioned for and enrolled in the choral program at ATHS. Therefore, students' acceptance of the existing assessment practices, as demonstrated by their continued participation in choir, shaped their perspectives and experiences.

Students who elect to participate in high school choirs may share demographic characteristics (Elpus & Abril, 2011) which are important to consider when interpreting findings relating to students' perceptions of fairness in assessment and grading practices. By selecting participants and sites for future studies that represent diverse demographics, communities, experiences, and grade levels, music educators can broaden their understanding of the complexities associated with assessment in choral music. Due to time constraints, I did not fully explore the demographic characteristics of the participants in this study or the implications of those factors; an interaction that is worthy of its own investigation.

Additionally, there could be alternative explanations for the students' positive perceptions of assessment in choir. Further exploration of the participants' achievement in other subjects in school and their motivation(s) for enrolling in choir would provide clarity and depth to the present findings. The voices of students who dropped out of choir, due to feeling alienated by assessment strategies, are absent from the present study. Their voices are critical to the discourse and continued study of assessment in choir.

Finally, I tended to conflate the terms assessment and grading throughout the study. Though my initial goal was to examine a variety of assessment practices (formal/informal, formative/summative), the participants frequently associated assessment with grading and this association informed their responses.

Conclusions and Implications for Music Education

Like the many investigations that preceded it, the findings of the present study shed light on some core difficulties that are associated with music assessment. For example, at ATHS, there were few administrative expectations that Ms. Andrews produce or submit documentation of individual student learning other than the summative grade that students received at the end of each grading period. She was, however, ethically compelled to implement assessments because they supported, in her mind and in the minds of her students, the goals that she developed for her classes. The learning outcomes associated with such goals generally aligned with state and national standards.

To create assessments in widely diverse classes, Ms. Andrews relied heavily on assess-

ments of nonmusical criteria—primarily participation and attendance. Ms. Andrews, her students, and their parents perceived such evaluations, though not specifically connected to music knowledge learning outcomes, to be appropriate and to support the ATHS music program goals, which were primarily associated with ensemble success and the cultivation of a life-long love of music. In short, I concluded that the participants' beliefs that nonmusical assessments were relevant to learning in this context superseded the various concerns about the validity and reliability of Ms. Andrews's nonmusical assessment practices.

Though limited to this study, the lived experiences, perceptions, and beliefs investigated here underscore the many tensions that exist in a profession that seeks to be a valued part of the academic community while honoring the large-group ensemble tradition rooted in our music education culture. Attempts to reconcile the nature of teaching and learning in choir with calls to adopt achievement criteria assessments that mirror individual assessment practices in other content areas further confounds this conflict. A core belief at ATHS was that choir participation was positively perceived to be more like an activity and less like an academic class. This important finding also merits additional study.

The findings of this study did not specifically call into question the value of individualized achievement criteria assessment practices or the importance of acquiring musical skills and knowledge. Instead, they provided a rationale for practitioners and researchers to develop, reexamine, and identify individualized assessment practices that are relevant to the types of skills and knowledge that choristers and teachers value. Specifically, the music education community would benefit from a more nuanced understanding of participation as a music-related (achievement) assessment criterion that reinforces skills intrinsic to ensemble performance readiness and then determine how to assess these skills in a fair and equitable way.

By attempting to reform the performance ensemble assessment model to one that more easily conforms to assessment norms in education, music educators should carefully consider the value of vocal development, ensemble performance skills (participation), and the notion of team effort that was both relevant and foundational to the students' experiences in choir, in this case. We must identify and examine choral music assessment models that value the participatory and collaborative nature of choir without marginalizing certain student populations. Perhaps there exists an opportunity for music educators to espouse the position of choir in the school community as a place of empowerment and identity, embracing and highlighting the strengths of our choral programs instead of focusing on their deficits.

References

Aitchison, R. (1993). Student perceptions and preferences in multiple criteria evaluation systems. *Contributions to Music Education, 20,* 45–56. https://www.jstor.org/stable/24127330

Bergee, M. J. (2003). Faculty interjudge reliability of music performance evaluation. *Journal of Research in Music Education, 51,* 137–150. https://doi.org/10.2307%2F3345847

Broomhead, P. (2001). Individual expressive performance: Its relationship to ensemble achievement, technical achievement, and musical background. *Journal of Research in Music Education, 49,* 71–84. https://www.jstor.org/stable/3345811

Colwell, R. (2008). Music assessment in an increasingly politicized, accountability-driven educational environment. In T. S. Brophy (Ed.), *Assessment in music education: Integrating curriculum, theory, and practice* (pp. 3–16). GIA Publications.

Conway, C., & Jeffers, T. (2004). Parent, student, and teacher perceptions of assessment procedures in beginning instrumental music. *Bulletin of the Council for Research in Music Education, 160,* 16–25. http://www.jstor.org/stable/40319215

Crochet, L. S., & Green, S. K. (2012). Examining progress across time with practical assessments in ensemble settings. *Music Educators Journal, 98*(3), 49–54. https://doi.org/10.1177%2F0027432111435276

Denis, J. (2016) Assessment in music: A practitioner introduction to assessing students. *Update: Applications of Research in Music Education, 36*(3), 20–28. https://doi.org/10.1177%2F8755123317741489

Elpus, K., & Abril, C. R. (2011). High school music ensemble students in the United States: A demographic profile. *Journal of Research in Music Education, 59,* 128–145. https://doi-org.umiss.idm.oclc.org/10.1177%2F0022429411405207

Ferm Almqvist, C., Vinge, J., Väkevä, L., & Zandén, O. (2017). Assessment as learning in music education: The risk of "criteria compliance" replacing "learning" in the Scandinavian countries. *Research Studies in Music Education, 39,* 3–18. http://journals.sagepub.com/doi/10.1177/1321103X16676649

Hawkins, J. A. (2018). *Secondary choral music educators' use of technology-assisted assessment tools* [Doctoral dissertation, University of Illinois at Urbana-Champaign]. IDEALS. http://hdl.handle.net/2142/100956.

Harrison, S. D., Lebler, D., Carey, G., Hitchcock, M., & O'Bryan, J. (2013). Making music or gaining grades? Assessment practices in tertiary music ensembles. *British Journal of Music Education, 30,* 27–42. http://journals.cambridge.org/abstract_S0265051712000253

Henry, M. L., & Demorest, S. M. (1994). Individual sight-singing achievement in successful choral ensembles. *Update: Applications of Research in Music Education, 13*(1), 4–8. https://doi.org/10.1177/875512339401300102

Heron, M. P. (2013). *It works in theory: A sequential and comprehensive approach to teaching music theory in the performance classroom.* Retrieved February 20, 2019, from http://itworksintheory.com

Kancianic, P. M. (2006). *Classroom assessment in United States high school band programs: Methods, purposes, and influence* (Publication No. 3222315) [Doctoral dissertation, University of Maryland]. ProQuest Dissertations & Theses database.

Kotora Jr., E. J. (2005). Assessment practices in the choral music classroom: A survey of Ohio high school choral music teachers and college choral methods professors. *Contributions to Music Education, 32*(2), 65–80. http://www.jstor.org/stable/24127154

LaCognata, J. P. (2011). Student assessment in the high school band ensemble class. In T. Brophy (Ed.), *The practice of assessment in music education: Frameworks, models, and designs* (pp. 227–238). GIA Publications.

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,* 168–183. https://doi.org/10.1177%2F0022429410369836

Lehman, P. R. (1968). *Tests and measurements in music.* (Vol. 4). Prentice–Hall.

Lehman, P. R. (2008). Getting down to basics. In T.S. Brophy (Ed.), *Assessment in music education: Integrating curriculum, theory, and practice* (pp. 17–28). GIA Publications.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach.* Sage Publications.

McClung, A. C. (1996). *A descriptive study of learning assessment and grading practices in the high school choral music performance classroom* (Product No. 9700217) [Doctoral dissertation, Florida State University]. ProQuest Dissertations and Theses database.

McCoy, C. W. (1988). An exploratory study of grading criteria among selected Ohio ensemble directors. *Contributions to Music Education, 15,* 15–19. http://www.jstor.org/stable/24127441

McCoy, C. W. (1991). Grading students in performing groups: A comparison of principals' recommendations with directors' practices. *Journal of Research in Music Education, 39,* 181–190. https://doi.org/10.2307%2F3344718

McMillan, J. H. (2018). *Classroom assessment: Principles and practice that enhance student learning and motivation.* Pearson.

McQuarrie, S. H., & Sherwin, R. G. (2013). Assessment in music education: Relationships between classroom practice and professional publication topics. *Research & Issues in Music Education, 11*(1), 1–15. https://files.eric.ed.gov/fulltext/EJ1015691.pdf

National Association for Music Education. (2016). *Student assessment using Model Cornerstone Assessments.* http://www.nafme.org/my-classroom/standards/mcas-information-on-taking-part-in-the-field-testing/

National Core Arts Standards. (2014). *National Core Arts Standards: Dance, Media Arts, Music, Theatre and Visual Arts.* https://www.nationalartsstandards.org

National Center for Education Statistics. (2018). *National Center for Education Statistics. U.S. Department of Education.* https://nces.ed.gov

Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice.* Sage Publications.

Russell, J. A. (2011). Assessment and case law: Implications for the grading practices of music educators. *Music Educators Journal, 97*(3), 35–39. https://doi.org/10.1177%2F0027432110392051

Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education, 58,* 37–54. https://doi.org/10.1177%2F0022429409360062

Ryan, C., & Costa-Giomi, E. (2004). Attractiveness bias in the evaluation of young pianists' performances. *Journal of Research in Music Education, 52,* 141–154. https://doi.org/10.2307%2F3345436

Stake, R. E. (1995). *The art of case study research.* Sage Publications.

Tracy, L. H. (2002). *Assessing individual students in the high school choral ensemble: Issues and practices* (Publication No. 3065486) [Doctoral dissertation, Florida State University]. ProQuest Dissertations & Theses database.

Wesolowski, B. C. (2020). "Classroometrics": The validity, reliability, and fairness of classroom music assessments. *Music Educators Journal, 106*(3), 29–37. https://doi.org/10.1177%2F0027432119894634