Assessment Practices of Secondary Music Teachers

Joshua A. Russell¹ and James R. Austin²

Abstract
The purpose of this study was to survey assessment and grading practices employed by secondary music teachers throughout the southwestern region of the United States. Three main research questions guided the study: (1) What types of school district frameworks and classroom contexts are secondary music teachers operating within? (2) Which specific assessment and grading practices are employed most commonly by secondary music teachers? (3) Do any contextual or individual difference variables influence secondary music teachers’ assessment and grading practices? Two mailings yielded 352 total usable questionnaires (36%) with a sampling error of ±5%. Participants reported that their school districts emphasized letter grades and that music course grades were equally weighted with other course grades in calculating student grade point averages and generating credit toward graduation. Yet, music teachers seldom received administrative guidance or altered assessment approaches due to standards-based curriculum adoption. Participants based grades on a combination of achievement and non-achievement criteria, with non-achievement criteria receiving greater weight in determining grades. Although instructional time, number of students taught, and number of concert performances prepared/given had no substantive relationship with assessment decisions, grading practices were influenced by teaching level and teaching specialization.

Keywords
music assessment, secondary music educators

We are in a time of dramatic change—if not revolution, then at least a major rebellion—in assessment.

—Shuler (1996, p. 88)

¹University of Hartford, CT, USA
²University of Colorado at Boulder, USA

Corresponding Author:
Joshua A. Russell, The Hartt School, University of Hartford, 200 Bloomfield Avenue, West Hartford, CT 06117
Email: jorussell@hartford.edu
Assessment is at the center of current efforts to improve education in the United States, and the act of assessing student learning is one of the most important responsibilities a teacher assumes (Miller, Linn, & Gronlund, 2009). Despite a wealth of research on educational assessment, major advances in testing and other forms of assessment, and greater availability of assessment resources, many teachers continue to develop their own assessment approaches in isolation—without regard to district curricula or assessment policies, and without considering the assessment practices employed by their colleagues or recommended by experts (Allen, 2005; Cizek, Fitzgerald, & Rachor, 1995/1996). The net effect is that there is little professional consensus as to what teachers should assess, how they should assess, or when they should assess. As Stiggins (2001) observes, “It is as if someone somewhere in the distant past decided that teachers would teach, and they would need to know nothing about accurate assessment” (p. 5).

A discussion of assessment in music education can spark lively and passionate debate. Many music educators, for example, share the concerns of other educators regarding the misuse of large-scale assessment data, lack of fairness in standardized testing, and the unintended consequences of high-stakes testing (Graham et al., 2002). Some music teachers believe the creative or interpretive nature of music precludes assessment but then readily employ subjective methods of assessment, many of which “are determined haphazardly, ritualistically, and/or with disregard for available objective information” (Boyle & Radocy, 1987, p. 2). The perception that assessment interferes with teaching and learning collides with the view that “assessment information is invaluable to the teacher, student, parents, school, and community for determining the effectiveness of the music instruction in their schools” (Asmus, 1999, p. 22). Beyond differing perspectives regarding assessment, job realities may curtail music educators’ efforts to assess student learning adequately. Shuler (1996), for example, pointed to lack of preservice or in-service training, limited student contact time, large student populations/class sizes, and insufficient resources for collecting, managing, and storing assessment artifacts.

**Music Assessment Research**

Most of the research related to music assessment involves surveys of elementary and secondary music teachers’ assessment and grading practices. Over the years, researchers have confirmed that assessment approaches adopted by music teachers in elementary general (Barkley, 2006; Carter, 1986; Nightingale-Abell, 1994; Talley, 2005), secondary choral (Kotora, 2005; McClung, 1996; McCoy, 1988; Tracy, 2002), and secondary instrumental (Hanzlik, 2001; Hill, 1999; Kancianic, 2006; McCoy, 1988; Sears, 2002; Sherman, 2006; Simanton, 2000) contexts are very idiosyncratic and not always well aligned with the recommendations of assessment experts.

Elementary general music teachers, for example, may view assessment as a means of adapting instruction and legitimizing general music as part of the elementary school curriculum (Talley, 2005). Yet, general music specialists have been found to favor informal assessment (observation, mental record keeping, subjective impression) over formal assessment techniques such as paper-and-pencil tests, projects, or structured
performance tasks (Barkley, 2006; Nightingale-Abell, 1994) and to grade based on participation and effort rather than actual student achievement in relation to objective standards (Barkley, 2006; Carter, 1986).

Researchers examining assessment practices in secondary music classrooms found that attendance and attitude were the most common grading criteria employed by instrumental and choral music teachers. Moreover, non-achievement criteria such as attendance, attitude, effort, and participation may be given more overall weight in the grading process than achievement criteria. McCoy (1991), for example, reported average weights of 42% for attendance, 33% for performance skill, 14% for attitude, and 11% for music knowledge. More recently, Simanton (2000) found that attendance (26% of grade) and attitude (30% of grade) accounted for more of the grade than performance (36% of grade) and written (8% of grade) assessments combined.

What factors might influence the assessment and grading practices of music teachers? McCoy (1988) suggested that guidance from school administrators was an important consideration. Principals placed greater emphasis on music knowledge and skill and less emphasis on attendance and attitude than did the high school ensemble directors surveyed by McCoy. Yet, inadequate administrative guidance is cited as a reason that music teachers do not demonstrate stronger assessment practices (Kotora, 2005; Nightingale-Abell, 1994).

McCoy (1988, 1991) also identified clear differences in how choral and instrumental music teachers selected and weighted grading criteria, implying that one’s teaching specialization within music may affect how students are assessed. Personal characteristics, such as teacher experience and training, have been associated with grading practices in some studies (McCoy, 1991; Simanton, 2000) but not in others (Hanzlik, 2001). There is some evidence that music teacher philosophy drives assessment decisions and priorities (Kancianic, 2006; Tracy, 2002).

Music teachers cite a range of situational factors that impede their efforts to implement new and improved assessment practices, including school size (Hanzlik, 2001; McCoy, 1991; Simanton, 2000), the large number of students being taught (Kancianic, 2006; Kotora, 2005; Lehman, 1998; Nightingale-Abell, 1994; Tracy, 2002), inadequate instructional time (Kotora, 2005; Nightingale-Abell, 1994; Tracy, 2002), difficulty in recording results and maintaining control of student behavior while conducting assessments (Kotora, 2005), parent and student apathy toward assessment in music classes (Kotora, 2005), and lack of training in assessment techniques (Kotora, 2005; Nightingale-Abell, 1994). Findings related to how such situational factors impact assessment and grading decisions, however, are inconclusive.

Music teachers have expressed the belief that assessment is an important responsibility (Hill, 1999) and that the types of grades awarded can impact the public’s perception of music’s value and status in the school curriculum (McClung, 1996). Grades also have been seen as a tool for motivating or disciplining students (Hill, 1999). While some music teachers may believe that grades should reflect specific learning objectives and individual music achievement (McClung, 1996), some may be satisfied with their current grading system whether or not it embodies recommended practices (Lehman, 1998; McClung, 1996; Simanton, 2000). This may be because grading
systems employed by music teachers (which typically produce a large percentage of “A” grades) tend to be endorsed by both students and parents (Hill, 1999).

Lehman (2000) viewed the adoption of national music education standards as a lever for making music assessment both necessary and of high quality. Yet, research does not reveal a discernable trend in the assessment of music students. While some individuals report revising their grading practices in recent years (Lehman, 1998) and are now including alternative forms of assessment such as portfolios in their grading scheme (Kotora, 2005), Barkley (2006) and Simanton (2000) showed that traditional assessment and grading approaches that emphasized non-achievement criteria and efficiency still dominated practice within the profession, and many music teachers continue to grade in ways similar to those reported by McCoy (1988) more than 20 years ago.

Study Purpose and Significance

The primary purpose of this study was to survey assessment and grading practices employed by secondary music teachers throughout the southwestern region of the United States.

In addition, we sought to resolve some contradictions evident in the literature by reexamining and identifying key factors that may influence how music teachers assess— including contextual factors as well as factors (education, experience, beliefs) that vary at the individual teacher level.

Three main research questions were addressed in this study: (1) What types of school district frameworks and classroom contexts are secondary music teachers operating within as they assess learning and grade students? (2) Which specific assessment and grading practices are most commonly employed by secondary music teachers? (3) Do any contextual or individual difference variables influence secondary music teachers’ assessment and grading practices?

This study is based on a pilot investigation (Austin, 2003) of high school band directors’ assessment and grading practices. In comparison to other state, regional, or national surveys related to music assessment, we attempted to more closely examine the specific assessment objectives, indicators, and formats underlying broad categories of grading criteria typically examined in prior research (attendance, attitude, knowledge, skill), and we further explored teacher–context–assessment connections. We also collected data from instrumental and choral music educators teaching at the middle school and high school levels in an effort to compare assessment and grading practices across teaching levels and specializations within music.

Methodology

Participants

The target population for this study was 4,889 secondary music teachers (orchestra, band, and choir directors) who were MENC members in the Southwestern division.
during the 2007–2008 school year. To ensure proportional representation of music teachers in each state, a stratified random sampling technique was employed \((n = 1000)\). An oversampling procedure also was utilized to reduce sampling error and increase generalizability. The initial mailing yielded 185 usable questionnaires, and a follow-up mailing yielded an additional 173 usable questionnaires (total of 352 usable questionnaires), resulting in an overall response rate of 36% with a sampling error of +/-5%.

Survey participants were highly representative of the target population with respect to state affiliation; a comparison of the number of secondary music teachers in the target population and participant sample within each state yielded an almost perfect correlation \((r = .991)\). The majority of participants (52%) indicated that band was their primary teaching assignment (37% taught choir, 11% taught orchestra), taught music in middle school/junior high (62%), and held a master’s degree (60%). Many participants reported prior assessment training, including conference clinics (87%), district in-service sessions (64%), graduate courses (57%), and university workshops (48%).

**Instrument**

The Secondary School Music Assessment Questionnaire (SSMAQ), adapted from a measure previously developed by Austin (2003), was the main research instrument used in this study. The questionnaire was organized into three main sections. Section One addressed the school context or assessment framework (school district policies, support, and influence). In Section Two, music teachers identified the specific assessment strategies and formats used in their classroom and provided additional information about grading criteria, grading weights, and grades typically awarded to their students. Items in Section Three focused on music teacher background (education, teaching experience, and assessment training) and classroom context (grade levels taught, instructional time with students, and typical ensemble performance responsibilities).

Development of the survey instrument began with a thorough review of questionnaires used in prior studies of secondary teachers’ assessment and grading practices (e.g., Cross & Frary, 1999; Hanzlik, 2001; McMillan, 2001; Simanton, 2000). A preliminary pool of questionnaire items was derived with two main considerations in mind: maximizing the content validity of items for secondary music ensemble class contexts (i.e., items taken from questionnaires targeting reading, math, or science were adapted to the extent possible) and optimizing response time while minimizing response errors (i.e., items requiring checklist or short-answer responses were favored over items requiring scale interpretation or open-ended responses). Through a piloting process involving 10 experienced secondary music teachers with assessment expertise, a draft version of the questionnaire was reduced from 40 items to 30 items. We eliminated items that provided less critical or redundant information, and we further modified the wording and response format for items that seemed confusing or difficult to answer. Estimated completion time for the SSMAQ was 8 minutes or less (certain sections of the questionnaire only were completed if corresponding assessment criteria were included in the music teachers’ grading policy).
**Procedures**

After obtaining campus Institutional Review Board approval, we sent the SSMAQ to a random sample of Southwest MENC members teaching in 2007–2008, including a self-addressed stamped envelope and a cover letter that detailed the study purpose, value of participation, and response deadline. Participants were informed that completing the anonymous questionnaire was voluntary and any information that may identify any participant would remain confidential. Two weeks after the initial deadline had passed, a follow-up questionnaire was mailed to the sampled music teachers. Given the large sample size, a significance level of .01 was established for all statistical analyses.

**Results**

**School District Assessment Framework**

In order to contextualize specific findings related to assessment and grading practices, we asked secondary music teachers to provide information about district-level policies/influences and the job realities they must address within the classroom. The vast majority of participants (95%) reported that their school districts award traditional letter grades alone. Apparently, few districts employed standards-based formats (2%), pass/fail formats (2%), or non-graded formats (1%) when reporting student achievement. In the majority of school districts (83%), ensemble grades were equally weighted with other academic subjects in determining student grade point averages. Similarly, 86% of districts awarded academic credit toward graduation for ensemble participation.

In addition to district-level policies related to grading, we examined how input from administrators and curricular developments (as contributors to the school assessment culture) may have influenced music teachers’ assessment practices. According to almost all study participants (92%), administrators seldom monitored or offered guidance as to how students enrolled in music ensemble classes should be assessed. Additionally, most music teachers (70%) reported that the adoption of standards-based curricula had little or no impact on their assessment practices.

**Classroom Context**

Beyond district assessment frameworks and the assessment culture that may permeate a particular school, we were interested in classroom-level factors that might impact secondary music teachers’ assessment practices. On average, participants reported having 224 min of instructional time per week with each ensemble. Some individuals, however, had as little as 45 min a week with their ensembles while others rehearsed their ensembles for as much as 460 min per week. With respect to teaching/assessment load, music teachers were responsible for assessing an average of 117 students (response range = 13–400) and preparing an average of 10 major ensemble performances each year (response range = 1–35).
Classroom Assessment Practices

The SSMAQ yielded data on a comprehensive set of classroom assessment practices, ranging from grading policies in the broadest sense to very specific information about assessment criteria, objectives, and formats. The majority of participants (79%) indicated that they provided students with a formal grading policy in writing—a practice that aligns with the recommendations of assessment experts. For one in five music teachers, however, information about the grading policy only was summarized verbally or not provided to students at all.

We also asked participants to identify the assessment criteria built into their grading system, the relative weight given to different criteria, and the percentage of ensemble students who receive various letter grades (see Table 1). Most music teachers used a combination or “hodgepodge” of achievement and non-achievement criteria to determine student grades. The greatest grade weights were allotted to performance, attitude, and attendance assessments. The least weighted assessments involved written knowledge and student practice. Grades assigned to music students were based primarily (60% of grade weight) on non-achievement criteria. The large variability corresponding to grade weights implies that there is little consensus as to the best way to calculate overall grades in secondary music classes. While no music teachers claimed to award all As, there was little evidence of grade differentiation. Participants indicated that the vast majority of students receive, on average, As (75%) or Bs (15%), while only 7% receive Cs and 3% receive Ds or Fs.

In addition to general approaches used to assign grade weights and grading outcomes, participants provided data about their specific assessment objectives and formats. In Table 2, results for teachers who include attendance and attitude as assessment criteria are summarized. When assessing attendance, music teachers typically considered student presence at major school concerts and daily rehearsals, as well as overall punctuality. Of the participants who based grades on attendance, 67% indicated that they reduced grades by one letter or more for unexcused absences from major events and daily rehearsals, while 30% implemented a partial grade reduction.

<table>
<thead>
<tr>
<th>Table 1. Assessment Criteria and Weights Used in Grading Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Achievement</td>
</tr>
<tr>
<td>Performance/skill</td>
</tr>
<tr>
<td>Knowledge</td>
</tr>
<tr>
<td>Nonachievement</td>
</tr>
<tr>
<td>Attendance</td>
</tr>
<tr>
<td>Attitude</td>
</tr>
<tr>
<td>Practice</td>
</tr>
</tbody>
</table>
Attitude most commonly was assessed based on in-class participation, responsibility (being prepared for class with all materials), effort, and citizenship. The majority of participants (71%) indicated that they used a combination of subjective impressions and objective documentation to assess attitudinal criteria. Others relied on completely objective documentation (10%) or completely subjective impressions (19%) to assess student attitude.

Music practice typically was assessed via paper reports or cards that documented the amount of time students practice (87%). Qualitative reports that describe what and how students practice were utilized by 30% of music teachers, while recordings of student practice sessions were assessed by roughly 1 in 10 music teachers. Because of the method of practice assessment employed by most teacher participants in this study (emphasizing quantitative or descriptive measures over evaluative measures), we classified practice assessment as a non-achievement grading criterion.

In addition to non-achievement criteria, we determined how achievement criteria (skills and knowledge) and associated assessment formats were factored into music student grades. The written assessment formats most commonly used by secondary music teachers were quizzes, worksheets, and major exams (see Table 3). Teachers were less inclined to use written assessment formats (homework, projects, and journals) that required out-of-class work by students and/or more extensive grading by the teacher. The majority of participants used written assessments to target basic knowledge of music terminology, symbols, and notation, as well as student ability to read or listen to music perceptively. Written formats also were used by a large proportion of music teachers to assess knowledge of music theory, music history, performance/pedagogical practice, and cultural contexts. Knowledge of compositional techniques and the ability to create small-scale compositions/arrangements, by contrast, were seldom assessed via written formats.

### Table 2. Factors Considered by Music Teachers When Assessing Attendance (N = 326) and Attitude (N = 332)

<table>
<thead>
<tr>
<th>Attendance factors</th>
<th>% of directors</th>
<th>Attitudinal factors</th>
<th>% of directors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major school concerts</td>
<td>97</td>
<td>In-class participation</td>
<td>96</td>
</tr>
<tr>
<td>Daily rehearsals</td>
<td>81</td>
<td>Responsibility</td>
<td>92</td>
</tr>
<tr>
<td>Punctuality</td>
<td>76</td>
<td>Effort</td>
<td>89</td>
</tr>
<tr>
<td>After-school rehearsals</td>
<td>42</td>
<td>Citizenship</td>
<td>81</td>
</tr>
<tr>
<td>Solo or large-group festivals</td>
<td>41</td>
<td>Instrument/uniform care</td>
<td>27</td>
</tr>
<tr>
<td>Athletic events</td>
<td>36</td>
<td>Leadership</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State festival participation</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Honor group participation</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ensemble support activities</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Private lesson participation</td>
<td>6</td>
</tr>
</tbody>
</table>

Attitude most commonly was assessed based on in-class participation, responsibility (being prepared for class with all materials), effort, and citizenship. The majority of participants (71%) indicated that they used a combination of subjective impressions and objective documentation to assess attitudinal criteria. Others relied on completely objective documentation (10%) or completely subjective impressions (19%) to assess student attitude.
The most commonly used formats for assessing students’ skills included playing exams (live, in class), concert performances, and sectional performances during class (see Table 4). Performance assessments typically addressed student technique and preparation of ensemble music as opposed to improvised performance, memorized performance, or sightreading. When scoring performance assessments, music teachers employed a variety of tools including rubrics (49%), rating scales (39%), and checklists (32%) or relied on their global impressions (i.e., everyday interactions and informal observations) (29%).

Influence of Contextual and Individual Difference Variables

Music teachers working in different kinds of contexts, managing different types of responsibilities, and possessing different attributes might be expected to adopt different approaches to assessment and grading. In an effort to better understand how varied approaches to assessment might become manifest in grading practices, we focused on the weighting of various grading criteria commonly employed by music teachers. Grade weights assigned to specific assessment criteria provide, arguably, the best evidence of a music teacher’s assessment philosophy and priorities.

Multivariate analysis of variance (MANOVA) was used to determine whether differences in grade weights might be attributed to school context variables such as instructional level (middle school/junior high, high school) and teaching specialization (instrumental, choral). For the multivariate analysis of school context effects, teaching level (middle school, high school) and teaching specialization (instrumental, choral)
served as the grouping variables, while the dependent variable set included grade weights corresponding to attendance, attitude, practice, knowledge, and skill (see Table 5). The multivariate analysis revealed a significant interaction effect for teaching level and teaching specialization ($\Lambda = 0.93, p < .001$) as well as significant main effects for teaching level ($\Lambda = 0.90, p < .001$) and teaching specialization ($\Lambda = 0.74, p < .001$). Follow-up univariate tests determined which mean differences in grading criteria weights contributed to the significant multivariate outcome.

Significant interactions existed for two dependent variables—knowledge ($F = 7.73, p < .01$) and practice ($F = 16.72, p < .001$). Simple main effects tests were used to clarify the nature of the interactions. Middle school choral directors gave significantly

### Table 4. Objectives and Formats for Performance Assessment ($N = 327$)

<table>
<thead>
<tr>
<th>Objectives</th>
<th>% of teachers</th>
<th>Formats</th>
<th>% of teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technique (scales, études)</td>
<td>67</td>
<td>Playing exam, live, in class</td>
<td>82</td>
</tr>
<tr>
<td>Prepared performance of ensemble music excerpts</td>
<td>64</td>
<td>Ensemble concert performance</td>
<td>52</td>
</tr>
<tr>
<td>Prepared performance of solo/chamber repertoire</td>
<td>46</td>
<td>Sectional performance in class</td>
<td>48</td>
</tr>
<tr>
<td>Sight reading</td>
<td>33</td>
<td>Auditions</td>
<td>34</td>
</tr>
<tr>
<td>Memorized performance</td>
<td>25</td>
<td>Playing exams, audiotaped</td>
<td>32</td>
</tr>
<tr>
<td>Improvised performance</td>
<td>7</td>
<td>Playing exams, live, out of class</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chair challenges</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Festival ratings</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Solo/ensemble festival ratings</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Playing exams, videotaped</td>
<td>6</td>
</tr>
</tbody>
</table>

### Table 5. Summary of Multivariate and Univariate Analyses for School Context ($N = 352$)

<table>
<thead>
<tr>
<th>Teaching level/specialization</th>
<th>Teaching level</th>
<th>Teaching specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda = 0.90^{**}$</td>
<td>$\Lambda = 0.74^{**}$</td>
<td>$\Lambda = 0.93^{**}$</td>
</tr>
<tr>
<td>Attendance</td>
<td>17.84^{**}</td>
<td>3.02</td>
</tr>
<tr>
<td>Attitude</td>
<td>5.19</td>
<td>60.17^{**}</td>
</tr>
<tr>
<td>Knowledge</td>
<td>5.30</td>
<td>14.44^{**}</td>
</tr>
<tr>
<td>Performance</td>
<td>0.001</td>
<td>23.37^{**}</td>
</tr>
<tr>
<td>Practice</td>
<td>13.93^{**}</td>
<td>29.88^{**}</td>
</tr>
</tbody>
</table>

Wilks's Lambda ($\Lambda$) values in header, $F$ values in cells.

* $p < .01$. ** $p < .001$. 
more weight ($M = 16.9\%$) to written assessment of music knowledge than middle school instrumental directors ($M = 9.6\%$), while there was no significant difference in the amount of weight given to music knowledge by high school instrumental and choral directors. Conversely, middle school band directors gave significantly more weight to practice assessments ($M = 11.9\%$) than did middle school choir directors ($M = 2.6\%$). As with knowledge assessment, however, there was no significant difference in the amount of weight given to practice by high school instrumental and choral directors.

Significant main effects, based on teaching level, emerged for attendance. High school directors gave greater weight to attendance ($M = 28.2\%$) than did middle school directors ($M = 18.5\%$). Significant main effects, based on teaching specialization, emerged for attitude and performance. Choral directors gave greater weight to attitude ($M = 37.5\%$) than did instrumental directors ($M = 21.0\%$), while instrumental directors gave greater weight to performance assessments of musical skill ($M = 31.5\%$) than did choral directors ($M = 21.2\%$). While numerous significant differences in mean weights assigned to various assessment/grading criteria emerged, it is important to note that all effect sizes ($\eta^2$) for the univariate analyses were less than .10, which suggests that these particular contextual variables (reflecting primary teaching positions or responsibilities) are accounting for only a small proportion of the variance in grade weights.

Correlational analysis was used to determine whether professional context variables (i.e., influence attributed to administrative guidance and adoption of standards-based curricula), classroom context variables (instructional time, number of students taught, number of ensemble performances), and individual difference variables (teaching experience, highest music education degree, and assessment confidence) might be associated with the pattern of grade weights assigned to assessment criteria. Several significant ($p < .01$) correlations of a weak to modest magnitude were found. Secondary music teachers who reported greater administrative guidance or oversight related to assessment and grading tended to give less weight to attitudinal criteria ($r = -.144$). Participants who reported having greater instructional time were less inclined to base grades on attitude ($r = -.153$), as were music teachers responsible for a greater number of performances ($r = -.196$). Moreover, those responsible for preparing more performances were less likely to emphasize student knowledge ($r = -.159$) and more likely to give greater weight to attendance ($r = .168$). Finally, participants who reported greater confidence in assessment were more likely to weight performance skills ($r = .231$) in music student grade calculations. The overall pattern of significant correlations (mean $|r| = .175$), however, suggests that grade weights (a proxy for assessment philosophy and priorities) are not related in any meaningful way to contextual or individual difference variables as measured in this study (see Table 6).

**Discussion**

The ubiquitous nature of assessment in American schools, coupled with the need for assessment strategies that contribute to learning and objectively document student
achievement, requires that teachers work diligently to enhance the process of educational assessment. If educational assessment trends and mores continue to lag behind assessment principles and policies as endorsed by experts, the courts may enforce increasingly stringent assessment requirements that demand a focus on student achievement (Dayton & Dupre, 2005). Music teachers, among all educators, would perhaps face the greatest challenge in moving their assessment paradigm out of the mid 20th century and into the 21st century. As Fiese and Fiese (2001) have observed, “music teachers need to be vigilant in making certain that they are planning to assess those behaviors that define the essence of the instructional goals” (p. 14).

Findings related to our first research question (i.e., dynamics of school district assessment framework and culture) suggest that established policies and common practice are at odds. Secondary music educators reported that their school districts emphasized traditional letter grades and that music course grades were equally weighted with other course grades in calculating student grade point averages and generating credit toward graduation. Yet, music teachers seldom received administrative guidance in assessment or changed their assessment approach in response to standards-based curriculum adoption. On one hand, policies related to grade point calculation and the awarding of graduation credit imply that student learning and methods of assessment matter just as much in music as in other so-called core subjects. On the other hand, secondary music teachers are given extraordinary autonomy and little support or guidance in relation to how they assess. It would appear that a “culture of benign neglect” exists—one that allows secondary music teachers to maintain status quo assessment practices without consequence.

Our second research question focused on the specific assessment and grading practices employed in secondary music ensemble classes. Secondary music teachers based grades on a combination of achievement and non-achievement criteria, with non-achievement criteria (attendance, attitude, quantitative measures of practice) receiving greater emphasis or weight in determining music student grades. This practice may

Table 6. Correlations Between Grade Weights and Contextual Variables (N = 329)

<table>
<thead>
<tr>
<th></th>
<th>Attendance</th>
<th>Attitude</th>
<th>Knowledge</th>
<th>Performance</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative guidance</td>
<td>-.077</td>
<td>-.144*</td>
<td>.140</td>
<td>.081</td>
<td>.101</td>
</tr>
<tr>
<td>Standards-based</td>
<td>-.115</td>
<td>-.078</td>
<td>.127</td>
<td>.052</td>
<td>-.045</td>
</tr>
<tr>
<td>curriculum adoption</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional time</td>
<td>.116</td>
<td>-.153*</td>
<td>-.029</td>
<td>.047</td>
<td>-.141</td>
</tr>
<tr>
<td>Number of</td>
<td>.107</td>
<td>-.095</td>
<td>-.033</td>
<td>-.026</td>
<td>.081</td>
</tr>
<tr>
<td>ensemble students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of performances</td>
<td>.168*</td>
<td>-.196*</td>
<td>-.159*</td>
<td>.112</td>
<td>-.017</td>
</tr>
<tr>
<td>Assessment confidence</td>
<td>-.071</td>
<td>-.051</td>
<td>-.034</td>
<td>.231*</td>
<td>.031</td>
</tr>
<tr>
<td>Teaching experience</td>
<td>.119</td>
<td>-.086</td>
<td>-.050</td>
<td>.022</td>
<td>.011</td>
</tr>
<tr>
<td>Highest earned degree</td>
<td>.013</td>
<td>-.023</td>
<td>-.043</td>
<td>.032</td>
<td>.040</td>
</tr>
</tbody>
</table>

*p < .01.
reflect the view, espoused by some music educators, that attendance and a positive attitude are prerequisites for student achievement. While some of the assessment objectives, formats, and practices utilized by music teachers were aligned with expert recommendations (e.g., development and dissemination of formal grading policies, use of written assessments to capture a wide range of music knowledge, frequent performance assessments, and varied tools used to increase reliability of performance assessment), other objectives, formats, and practices would hardly be considered assessment exemplars (e.g., giving attendance extensive weight in the grade formulation and issuing substantial grade reductions on the basis of attendance, relying on subjective opinion to assess student attitude, emphasizing quantitative measures of practice, neglecting assessment in the creative domain, emphasizing prepared performance of ensemble repertoire rather than performance indicators of musical independence and learning transfer, and awarding a very large proportion of high grades). Given the overall status of assessment and grading practices in secondary music, reducing music teacher reliance on behavioral or managerial aspects of grading, rather than eliminating such practices, may be a realistic and attainable goal for the profession (Johnson, 2008).

Our perceived lack of sophistication and progress in how secondary music teachers assess learning and grade students may reflect the very nature of music as a discipline and music teachers as a culture unto themselves (Morrison, 2001). Secondary music classes typically are elective in nature, and there is evidence that teachers of elective subjects, as compared with other teachers, occasionally award higher grades that are based on more subjective, non-achievement criteria (Bridgham, 1973; Dietrich, 1973; Johnson, 2008), particularly individuals who are concerned about maintaining adequate course enrollments (to protect courses or even teaching positions). Music instructors also must cope with the reality of addressing both curricular and extracurricular responsibilities (e.g., performances at school concerts, state-sanctioned festivals, school/community functions), and these additional pressures might play a role in making assessment in music classes distinct from methods found in traditional academic classes. Finally, music itself is an artistic process, and as such, music educators might be expected to hold different views about the nature, purpose, and importance of assessment. There is evidence of subject-area differences in how secondary teachers grade (Duncan & Noonan, 2007; McMillan, 2001), with teachers of performing and fine arts subjects, for example, tending to emphasize enabling behaviors (attendance, participation, effort) to a greater extent than teachers of so-called core subjects (English, math, science, social studies).

Our third research question pertained to the influence of contextual and individual difference factors on assessment and grading practices. Many music teachers and leaders within the profession are inclined to attribute subpar or outmoded assessment to external factors that are seemingly beyond the control of individual educators—impossible workloads, performance pressures, not enough time, and not enough training or professional development. We found little evidence, however, that assessment decisions and practices reported by secondary music teachers reflect untenable situations.
The majority of participants in our study appear to work under adequate, if not ideal, classroom conditions. Moreover, issues of instructional time and number of students taught had no substantive relationship with assessment decisions or grading priorities (i.e., weight assigned to achievement vs. non-achievement assessment criteria). Some music teachers who were responsible for very busy schedules and many students, for example, were among the most sophisticated in their choice of assessment strategies and the most credible in how they graded students.

Music teachers’ assessment and grading practices were influenced, to a certain extent, by teaching level and teaching specialization. Middle school choral teachers tended to give greater weight to knowledge, perhaps because of the role that understanding and interpreting text plays in choral performance, while middle school instrumental teachers were more inclined to give weight to practice assessments (practice reports are commonly linked to private or pullout lessons, experiences germane to instrumental music but rare in choral music). High school teachers assigned greater weight to attendance, perhaps because of a greater number of scheduled performance events, and less weight to attitude than middle school teachers, who perhaps are coping with and accounting for dramatic shifts in attitude among early adolescents. Instrumental music teachers assigned less weight to attitude assessments and greater weight to performance assessments than choral teachers. These differences may indicate that performance skills are either more valued or considered easier to assess in instrumental contexts. Alternatively, choral music teachers may emphasize attitudinal assessments, despite the challenges inherent in documenting and reliably assessing attitude, because of a stronger desire to cultivate social goals and sense of community than instrumental teachers (Kennedy, 2003).

Administrative guidance in assessment was significantly associated with music teachers’ assessment of attitude; greater guidance was associated with less grade weight. Most administrators likely are aware of the subjectivity and legal risks that surround attitudinal assessment and would therefore discourage attempts to base grades extensively on attitude. Additionally, secondary music teachers who reported being more confident about their assessment abilities tended to accentuate performance assessments when determining grades. There is a strong tradition of performance assessment in secondary ensemble classes, and familiarity with this particular facet of assessment probably contributes to greater overall confidence in assessing music students. Positive and near-significant relationships also emerged between music teachers’ weighting of music knowledge and influence attributed to both administrative guidance and standards-based curriculum adoption. Collectively, these findings suggest that the school assessment culture, in consort with individual teacher training to the extent that it yields increased confidence, may serve to elevate assessment practices adopted by secondary music teachers (Sato, Wei, & Darling-Hammond, 2008).

We launched this study with the intent of identifying some key interventions that might be targeted so as to improve the art and science of music assessment. Based on our findings, we offer a few suggestions. First, in an effort to emphasize
achievement-based assessment and deemphasize the use of attendance and/or attitude to determine student grades in music, standards-based curricula should be considered a “point of departure” in formulating assessment strategies. Prior research suggests that the National Standards are not a primary concern in assessment (Kancianic, 2006), and we did not find any statistically significant relationships between standards-based curriculum adoption influence and typical weights assigned to grading criteria. Nonetheless, music teachers are not likely to emphasize written assessment of music knowledge on par with other grading criteria until they recognize and accept the Standards as a framework for assessing students. As Shuler (1996) has indicated, the National Standards for Music Education give music educators a more elaborate learning palate from which to assess.

Because many school administrators do not appear to support or greatly influence the assessment process within secondary music contexts, it is critically important that music teachers share, discuss, and evaluate assessment strategies with their colleagues, provided that any discussion or evaluation focuses on the relative effectiveness of those various strategies rather than efficiency alone. Beyond collegial exchanges and collaborations, there is continued need for general assessment training (Brookhart, 2001; McMunn, Schenck, & McColskey, 2003; Mertler, 2004). This training may come in the form of a required course or common unit on assessment for all preservice music teachers and sustained professional development for music teachers in the field. Assessment training should engage music teachers in rethinking their beliefs about the fundamental purpose of assessment (Mertler, 2004), considering the underlying ethical and legal ramifications of their classroom assessment practices (Green, Johnson, Kim, & Pope, 2006), and exploring how assessment might best be used to improve the learning of individual students in music contexts. We did observe a relationship (likely reciprocal) between music teacher confidence in carrying out assessment responsibilities and decisions they made related to weighting performance assessment criteria. Confidence, as a psychological construct, may represent a critical bridge between assessment training and the actual implementation of that training in developing more comprehensive and valid assessment frameworks within the music classroom context.

More research should be conducted to determine the test–retest reliability of the SSMAQ as well as the reliability of assessment techniques actually employed by secondary music teachers. In addition, research studies emphasizing different facets of assessment (i.e., assessment values and beliefs) and/or employing different methodologies (projective techniques, content analyses, case methods) may prove useful in identifying assessment models/exemplars while further increasing practitioners’ understanding of various factors that may influence their day-to-day decisions related to assessment and grading.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the authorship and/or publication of this article.
Financial Disclosure/Funding

The authors received no financial support for the research and/or authorship of this article.

References


**Bios**

**Joshua A. Russell** is assistant professor of music education at The Hartt School of the University of Hartford. His research interests include education policy, teacher education, string education, and psychosocial/cognitive development in music learning and teaching.

**James R. Austin** is professor of music and associate dean of undergraduate studies at the University of Colorado at Boulder. His research interests include student motivation and self-concept development, teacher education, classroom-level assessment, and school/arts policy implications of educational reform.

Submitted May 25, 2009; accepted October 22, 2009.