2017

An investigation of the assessment practices of elementary general music teachers in Iowa

Michelle Linn Hyde Swanson
University of Northern Iowa

AN INVESTIGATION OF THE ASSESSMENT PRACTICES OF ELEMENTARY GENERAL MUSIC TEACHERS IN IOWA

An Abstract of a Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

_____

Dr. Robert Boody, Committee Chair

_____

Dr. Kavita Dhanwada
Dean of the Graduate College

Michelle Linn Hyde Swanson

University of Northern Iowa

July, 2017

ABSTRACT

Assessment, leading to informed decisions, is an integral part of music education and may be a means to justify the importance of music as a valued content subject (Brophy, 2000). The purpose of this study was to describe the elementary general music assessment practices in Iowa. The study was framed using Stiggins' (2005) categorical definition of assessment as gathering, organizing, summarizing, and reporting data. Mixed methodologies were utilized to gather data regarding these four categories through survey questions (N=211) and two focus group discussions. The participants were from a variety of school district sizes as well as teachers with varying years of teaching experience,

Through the survey and focus groups, data were collected and analyzed to describe the (a) characteristics of data gathering, organizing, summarizing and reporting assessment practices; (b) demographics or teacher characteristics that relate to assessment practices; and (c) beliefs or opinions related to assessment in elementary general music in Iowa.

Findings from this study indicate that observation was the most prevalent means of gathering assessment data. Organizing and summarizing practices were variable, with no common standard reported by respondents. Both academic content and non-academic behavioral aspects were summarized by a majority of the respondents. Report cards were the most frequently used reporting tools with a majority of the respondents required to attend parent-teacher conferences. Statistical significance was found between certain demographic variables and other reported practices. As the school size of the respondents

increased, so did the reported level of challenges with assessment. As the years of teaching experience increased, so did the overall quantity of organizing and summarizing practices as well as the reported abilities in completing assessments in general. Attitudes toward assessment, in general, were negative, with barriers to assessment reported as large class sizes, demanding teaching schedules, large total numbers of students, and lack of preparation time. The focus groups also identified student misbehaviors as a barrier to gathering assessment data. Results from this study add to the literature on how elementary general music teachers (in Iowa) assess students.

AN INVESTIGATION OF THE ASSESSMENT PRACTICES OF ELEMENTARY GENERAL MUSIC TEACHERS IN IOWA

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

_____

Dr. Robert Boody, Chair

_____

Dr. Deborah Tidwell, Committee Member

_____

Dr. Linda Fitzgerald, Committee Member

_____

Dr. Cayla Bellamy, Committee Member

Michelle Linn Hyde Swanson

University of Northern Iowa

July, 2017

DEDICATION

This dissertation is dedicated to my husband, Chad; my children, Matthew and Elizabeth; and my parents, Dexter and Sandra Hyde. I love you.

IN MEMORIAM

This dissertation is in loving memory of my Aunt Beverly Riess who I miss so very much. She always encouraged me in my teaching, research, and creativity. I finished this for her. Also in memory of Dr. Lynn Nielsen who was an original member of my dissertation committee.

ACKNOWLEDGEMENTS

I would like to give my utmost thanks to my Lord and Savior, Jesus Christ. It is only by the grace of God that this degree and this research were ever completed. The signs were so clear and so timely. Thank you, God, for putting everything in place for this to happen successfully. May this work bring glory to Your name!

I would like to thank my family. My amazing husband has had to put up with an absentee and stressed out wife for years! Your patience and encouragement was overwhelming. Thanks for paying my u-bill and encouraging me to go back each semester! My children have been SO patient with me! No more grouchy-ouchy mom, I promise! I hope you value education and have seen how education has changed my life. I hope you always reach for *your* educational goals.

My parents have been unbelievable help. Between picking up my kids, doing laundry, dishes, running errands, fixing things... I never, ever would have finished this without you both. How many times did you drop me off at the library, so I wouldn't have to walk so far??? My sister also kept me able to get this done with her helpfulness, advice, and words of encouragement, and especially talking to me on the phone as I walked home from the library in the dark! Anne, our babysitter and pseudo-family member, also provided much support and cared for my kids during doc school!

I would like to thank all of my dissertation committee members for their patience and care with my topic and research. Without you, this would not have happened. I would very much like to thank my two wonderful advisors, Dr. Tidwell and Dr. Schneider. They were a blessing so many times throughout my studies. Their patience, advice, editing, and just uplifting encouragement was priceless. I have also had some

amazing instructors in my doctoral studies. Thank you for sharing your knowledge with me. Dr. Jeanne Steele, Dr. Charlotte Haselhuhn, Dr. Robert Boody, and Dr. Sandra Alper were some of my favorites!

Honestly, without the continual support and encouragement from Dr. Boody, I probably never would have finished. Thank you for never giving up on me.

I would also like to thank the staff at the UNI ROD Library. The circulation desk was very familiar with my face and their courteous and helpful service was invaluable. The Interlibrary Loan department, reference librarians, and the student helpers were unbelievably helpful and kind.

I also need to thank my wonderful music education students at UNI. Thank you for your patience as I continued to work, research, and write. The prayers from CPAF, the patience of my advisees, the specific support from Emily, Alyssa, Acacia, Kelly, Bonnie, and Ashley were so important.

To my church family, thank you for the prayers. God is so good! I am so blessed by your continued prayerful support and love.

Lastly, I would like to thank my wonderful friends. Without you, this never would have happened. Your continued prayerful support over the years was preciously important. Dyan, I never would have ever been able to conquer this degree without you, my sister in Christ, in my life. Joy, Nicole, Traci, Lynnette, Karen, Nichole, Brandy, and so many more - thank you. Thanks, too, for putting up with me NOT being available to do all of the things with you. No more studying for me! Let’s do lunch - every day!!!

TABLE OF CONTENTS

PAGE

LIST OF TABLES xi

CHAPTER 1. INTRODUCTION 1

Background to the Problem 1

Problem Statement 4

Purpose of Study 5

Research Questions 6

Significance of Study 7

Organization of Study 8

CHAPTER 2. REVIEW OF LITERATURE 10

History of Assessment 11

History of Assessment in Education 11
Changes in the twentieth century 13
Questionable practices 14
Impact on learners 15
Outside influences 15
Recent studies 17

History of Assessment in Music Education 20
Lack of credibility 21
Standardized tests in music education 22
National Standards in Music Education 25
Teachers' implementation of the National Standards in Music Education 29

Practices in Gathering Assessment Data 31

Specific Types of Data Gathering Tools 33
Examinations 33
Projects 34
Assignments and homework 34
Observations 34
Performances 36

Portfolios....37
Rubrics....39

Frequencies of Gathering Practices in Music....40

Practices in Organizing Assessment Data....41

Problematic Organizational Issues....42
Large number of students....42
Instructional time....43
Limited training....44
Measurement....47
Preparation time....47

Suggestions for Improvement in Organization....48
Efficiency....48
Organization....49
Time management....50

Practices in Summarizing Assessment Data....52

Academic Factors....53
Indicators and standards....53
Content....54
Categories....56

Nonacademic Factors....58
Participation....61
Attitude....62
Effort....63
Behavior....64
Punctuality....65

Practices in Reporting Assessment Data....66

Grades....67
Report cards....67
Conferences....70

Teachers' Motivations for Assessment....71

Teachers' Practices with Assessment in Music Education....77

Elementary Music Research Studies....78
Secondary Music Research Studies....83
Summary of Research Studies in Music Assessment....84

Summary 84

CHAPTER 3. METHODOLOGY 87

Rationale and Design 88

Rationale 88
Design 90

Population 90

Quantitative Study: Survey 91

Procedure for Selection of the Quantitative Sample 91

Development of Instrumentation 93
Pilot study 93
Literature review 95
Other surveys 98
Research expert 98

Description of Survey Instrument 98

Procedure 103

Analysis 104

Qualitative Study 104

Interviews 105
Criteria for interview sample selection 106
Data 106

Optional Review of Documents 107

Conclusion 108

CHAPTER 4. RESULTS 109

Quantitative Study: Survey 109

Demographics 110
Full-Time equivalency and number of students 115

Research Questions 116
Question 1A: Practices of gathering assessment data 117
Question 1B: Practices of organizing assessment data 124
Question 1C: Practices of summarizing assessment data 126
Question 1D: Practices of reporting assessment data 132
Question 2: Demographic comparisons with responses 140
School size 140
Number of students 142
Number of buildings 143
Teaching experience 144
Question 3: Teacher beliefs related to assessment practices 145
Demographics and opinions 151

Qualitative Study: Focus Groups 155

Small School District Summary 156
Gathering assessment data 157
Organizing assessment data 157
Summarizing assessment data 157
Reporting assessment data 158
Student behavior impact on assessment 158
Teacher beliefs regarding assessment 158

Large School District Summary 159
Gathering assessment data 159
Organizing assessment data 160
Summarizing assessment data 161
Reporting assessment data 161
Student behavior impact on assessment 161
Teacher beliefs regarding assessment 163

Summaries of Quantitative and Qualitative Data 164

Demographics 164

Research Question 1: Gathering, Organizing, Summarizing, & Reporting 165

Research Question 2: Demographic Implications 168

Research Question 3: Related Opinions 168

Conclusion 169

CHAPTER 5. CONCLUSIONS 171

Limitations 171

Delimitations 172

Data Summaries and Discussion 172

Demographics of Study Respondents and Participants 173

Research Question 1 174
Assessment gathering practices 175
Assessment organizational practices… 177
Assessment summarizing practices 178
Assessment reporting practices 179

Research Question 2 181

Research Question 3… 184
Quantitative and qualitative data 185

Relationship with Previous Research 188

Research Question 1 189

Research Question 2 189

Research Question 3 190

Implications and Recommendations 190

Suggestions for Future Research 192

Conclusion 193

REFERENCES 195

APPENDIX A: PILOT SURVEY 222

APPENDIX B: PRE-NOTICE LETTER 224

APPENDIX C: COVER LETTER 225

APPENDIX D: FIRST REMINDER LETTER 227

APPENDIX E: FINAL REMINDER LETTER.....228

APPENDIX F: SURVEY INSTRUMENT.....229

APPENDIX G: FOCUS GROUP PROTOCOL 239

APPENDIX H: INFORMED CONSENT FOR FOCUS GROUPS 240

APPENDIX I: DOCUMENT REVIEW CHECKLIST 242

APPENDIX J: MAP OF AERA EDUCATION AGENCY REGIONS 243

APPENDIX K: PHONE INVITATION TO PARTICIPATE: SMALL DISTRICTS ..244

APPENDIX L: PHONE INVITATION TO PARTICIPATE: LARGE DISTRICT 245

APPENDIX M: SMALL SCHOOL FOCUS GROUP CODING OF THEMES 246

APPENDIX N: LARGE SCHOOL FOCUS GROUP CODING OF THEMES 250

APPENDIX O: FOCUS GROUP TRANSCRIPT: SMALL DISTRICTS 255

APPENDIX P: FOCUS GROUP TRANSCRIPT: LARGE DISTRICT 264

LIST OF TABLES

TABLE PAGE

1 Research Studies in Music Teachers' Practices with Assessment....74
2 Sources of Survey Instrument Questions....100
3 Demographic Descriptors of Respondents: School Size and Years of Teaching..111
4 Representation of Respondents by Area Education Agency Region....112
5 Demographic Descriptors of Respondents: Degree, FTE, and Buildings 113
6 Teaching Variables of Respondents....115
7 Students Taught Per Week by Teaching Load....116
8 Frequencies of Respondents' Assessment Tools....118
9 Assessment Tool Use by Music Objective 119
10 Commercial Test Use of Respondents....121
11 Frequencies of Formal Assessments by Grade Level....122
12 Types of Music Objectives Assessed....123
13 Respondents' Practices in Organizing Assessment Data....124
14 Respondents' Reported Tendencies with Organizing Assessment Data 125
15 Respondent Opinions Regarding Assessment Summaries 127
16 Opinions on Behavioral Factors of Assessment Summaries 130
17 Two Questions on Behavioral Factors of Assessment Summaries....131
18 Practices Relating to Reporting Data....132
19 Frequency of Using Reporting Tools....134
20 Reported Categorical Options for Report Card Use 136

21 Report of District-Wide Mandated Report Cards 137

22 Reported Changes in Report Card by Time or Grade.....138

23 Reported Requirements During Conferences 139

24 Crosstabs of Demographic Variables and Assessment Practices 142

25 Beliefs Regarding Educational Reasons for Assessments.....145

26 Teacher Responses of Requirements or Motivators for Assessment.....147

27 Respondents' Reported Greatest Challenges with Assessment.....148

28 Respondents' Reported Greatest Advantages to Assessment.....148

29 Respondents' Level of Satisfaction with Assessment 149

30 Participants' Opinions Regarding Assessment.....150

31 Crosstabs of Demographic Variables and Opinions 152

32 Variables Cross-Tabbed with Reported Opinions 154

33 Focus Group Coding Categories and Frequencies.....156

CHAPTER 1

INTRODUCTION

All teachers have many responsibilities, one of which is to assess students' academic achievement in the classroom (Allen & Lambating, 2001; Chiodo, 2001; Duncan, 2009; Goolsby, 1999; Hepworth-Osiowy, 2004; Sherman, 2006). Music teachers are not exempt from the expectation of evaluating students and describing student progress. Unfortunately, despite the National Standards in Music Education, there are few common, clear expectations, criteria, or guidelines for exactly how to evaluate, grade, and report musical growth (Brophy, 2003; Burbridge, 2001; Jindrich, 1996; Music Educators National Conference [MENC], 1996a, 1996b; Nightingale-Abell, 1993; Pontious, 2001; Walker, 2001; Wells, 2001). The literature review (Chapter 2) of elementary general music education assessment indicated a lack of research on this topic. This study will investigate the current assessment practices of elementary general music teachers in Iowa.

Background to the Problem

Improvement in education is often centered on assessment (Hoolsema, 2010; Miller, Linn & Gronlund, 2009; Russell & Austin, 2010). Assessment practices provide "information to students, parents, and teachers about students' progress" (Duncan, 2009, p. 1). Assessment protocol and practices continue to change, yet lack a consensus of what to assess, how to assess, or when to assess students (Farrell, 1997; Russell & Austin, 2010).

Assessment is one part of the overall teaching and learning experience and should be embedded within the goals and objectives of instruction of both general education and music education (Brophy, 2000; Lavender, 2000; Marzano, 2000; Shuler, 1996b; Stiggins, 1999; Wells, 2001). One of the many challenges to elementary general music teachers is to gather accurate information and then report that information to parents (Stiggins, 2005). Cope (1996) described the importance of assessment in music education since the *Goals 2000: Educate America Act* in 1994, and the "critical" (p. 39) necessity for music educators to excel in assessment techniques. Hepworth-Osiowy (2004) stated that music teachers are aware of the benefits of assessment in music education, but such assessment is considered one of the "greatest challenges" (p. 1) for music teachers.

Many authors have agreed with the challenge of assessment in music. Hoolsema (2010) stated that music teachers find assessment procedures difficult. Brophy (2003) stated even more strongly that many states lack assessment tools in music altogether. Music educators face many problematic situations with assessment due to limited training in assessment, minimal contact time with students, large numbers of music students, and minimal resources to accomplish assessment goals (Barkley, 2006; Brummett & Haywood, 1997; Hanzlik, 2001; Shuler, 1996b, Simanton, 2000). In 1974, Colwell attempted to determine why more music teachers were not often performing assessment tasks. Colwell reported several issues for not assessing: fear of exposing low quality teaching, avoidance of personal choices, lack of familiarity with testing, and lack of measurement principles.

Hoolsema (2010) described the current assessment scenario as “monitoring student achievement” (p. 19) and predicted a future mandate of music teachers to report “yearly data demonstrating student achievement and growth” (p. 19). With continued emphasis on high expectations and accountability, Farrell (1997) described not only a continuous need for evidence of quality, but also a focus on the measure of that growth toward excellence.

This aim toward excellence in music became a critical nation-wide issue in the early 1990s. The national music education advocacy organization at that time was known as Music Educators’ National Conference (MENC). The organization is now called The National Association of Music Educators (NAfME). Then MENC implemented national standards in music in 1994 and published numerous publications to serve as resources for implementing and assessing the standards (MENC, 1994a; 1994b; 1994c; 1994d; 1996a; 1996b; Lindeman, 2003). Beginning in the mid-nineties, MENC, through the standards, directed music teachers to focus assessments on “measurable objectives” (Duncan, 2009, p. 1).

In addition to the national emphasis on assessment, was also the overall documented dissatisfaction with assessment itself and a commitment to discover what is being used, whether it is working, and how it could be improved (Brummett, 1993; Carter, 1984; Chiodo, 2001; McClung, 1996; Niebur, 1997; Sherman, 2006). In music especially, teachers are seeking concrete, legitimate findings that demonstrate high quality learning and legitimize music education (McClung, 1996; Talley, 2005). As with all areas of education, music teachers are responsible for the quality of music instruction

that they deliver to children in music classes. Consistency through quality assessments in music could provide music teachers with the "evidence they need to report progress to parents and to satisfy accountability requirements" (Brophy, 2003, p. 15).

This general background to assessment in music education is significant for this study. As noted in the literature, discontent and dissatisfaction in areas relating to music assessment are problematic. Furthermore, a lack of consistency – nationwide – in music education assessment is dubious. Interested parties continue to look toward assessment in music education as a type of remedy, yet the assessment paradigm is questionable.

Problem Statement

Arguments have been made that knowledge of, information about, and familiarity with assessment are lacking and have led to many problematic situations (Cizek, 1996). Although research has been conducted specifically on assessment in music, many teachers continue to find challenges with assessment practices (Hepworth-Osiowy, 2004; Kotora, 2005; Marzano, 2000; McCoy, 1988; Monroe, 1995; Niebur, 1997; Nightingale-Abell, 1993; Russell & Austin, 2010). Since the inclusion of the 1994 National Standards in Music, further problems have arisen as many music educators have struggled with standards-based assessment, either not understanding the standards in general or not possessing the knowledge of how to base assessment around the standards. (Brophy, 2003). Studies have, furthermore, found while teachers have realized the shortcomings of their current assessment practices, few have found the means to grow in their assessment techniques or develop a reporting format that is satisfactory to the parties involved (Guskey, 2001; Hanzlik, 2001; Monroe, 1995).

If assessment is to play a part of an educational accountability solution, then assessment practices need to be clear, consistent over time, and transparent. The literature, however, did not describe such a scenario in general education or music education. Specifically in Iowa, as with music education nationwide, there is a lack of information or data describing the assessment practices in elementary general music education. Wesolowski (2012) stated, "Now more than ever, teachers… find themselves in situations which a thorough documentation of student performance is necessary" (p. 36). Is such a thorough documentation occurring in elementary general music education in Iowa?

Purpose of Study

The purpose of this study was to describe the current practices in assessment of elementary general music teachers in Iowa. Stiggins (2005) summarized classroom assessment into four categories: gathering evidence, organizing or storing evidence, summarizing evidence, and sharing or reporting evidence. Gathering evidence of student learning was described as the collection of data with regard to student performance or behavior with relation to learning objectives, through observation, work product review, or testing tasks. Organizing or storing evidence was described as a systemized collection process where each student's data were assembled in an orderly and timely fashion. Summarizing evidence was described as a synthesizing process of compiling raw scores or observational notes into meaningful, reflective terms with relation to achieving the learning objectives. I added further questions, for this study, to discover if Iowa music educators were including non-academic data during the summarizing process. Reporting

the evidence was described as communicating the student learning to other interested parties.

This study investigated the current practices of music teachers in Iowa using the four categories of gathering, organizing, summarizing and reporting, as described above by Stiggins (2005). If assessment is as vital and necessary as described in the literature, this study was designed to describe what is happening in elementary general music assessment in Iowa. Data regarding the four categories were gathered initially through survey questions, with two follow-up focus groups.

Research Questions

Through survey questions and interviews of Iowa music educators, practices will be examined to determine factors relating to Stiggins' categories as well as demographic connections to the categories as well as opinions regarding the categories. Specifically, the following research questions were addressed.

1. What are the features of the classroom assessment practices currently implemented to:
    a. gather evidence (assessment data) of student learning in elementary general music in Iowa?
    b. organize student learning data in elementary general music in Iowa?
    c. summarize assessment data in elementary general music in Iowa?
    d. report assessment data in elementary general music in Iowa?

2. What demographics or teacher characteristics influence or relate to classrooms assessment practices of gathering, organizing, summarizing, and reporting assessment data in elementary general music in Iowa?
3. What are teacher beliefs or opinions related to assessment and reporting in elementary general music in Iowa?

Significance of Study

As long as music teachers have been teaching, assessments of students' musical behaviors have occurred (Lopez, 2001). A review of the literature indicated, however, that little is known about specific assessment practices in elementary general music. Although thirteen studies have been published on elementary general music assessment practices, none specifically investigated elementary general music assessment in Iowa.

Assessment based on informed decisions is an integral part of music education (Brophy, 2000). Furthermore, in a time of national and state economic crises, confounded with an accountability-laden educational atmosphere, assessment may also provide a means to justify the importance of maintaining music as a valued content subject within school curricula. In addition, high quality assessments are beneficial to students and parents. Using assessments in music classes can aid students in future decisions with regard to music (Brophy, 2000).

> Conscientious and deliberate use of regular assessment can strengthen any program and provide valuable assistance to the student developing the skills needed to form a lifelong involvement with music (Cope, 1996, p. 40).

This study of assessment and reporting in elementary music is important and timely because of the perceived lack of a common system of music assessment data

gathering, organizing, summarizing, and reporting in Iowa, the four specified areas of assessment noted by Stiggins (2005). Davidson (1995) reported an overall difficulty for all music teachers in the "transition" (Davidson, 1995, p. 67) from a simple awareness of assessment issues to actual change in practices or strategies in assessment.

The main outcome of this study provided music educators with an overall summary of what was taking place in elementary general music education assessment throughout Iowa. A second outcome determined any relative demographics or characteristics among Iowa educators that influenced elementary general music assessment. Further outcomes described the opinions and beliefs related to elementary general music assessment in Iowa, and also recommended improved assessment practices. Results from this study added to the overall literature on how elementary general music teachers in Iowa were assessing students.

Organization of Study

The remainder of the dissertation was divided into four chapters. In Chapter 2, the literature review had the following organizational categories: history of assessment, practices in gathering assessment data, practices in organizing assessment data, practices in summarizing assessment data, practices in reporting assessment data, teachers' motivations for assessment, and teachers' practices with assessment data. In Chapter 3, the methodology was explained, which included explanations of the rationale and design of the study, definitions of the population and sample, and descriptions of the quantitative and qualitative research methods. In Chapter 4, the results of the data collection were

described. Finally, in Chapter 5, the data was interpreted, findings provided, and recommendations made.

CHAPTER 2

REVIEW OF LITERATURE

The purpose of this study was to examine the assessment practices in elementary general music classrooms in Iowa and music teachers' attitudes toward those practices. Throughout this study, teachers' practices in assessment was focused on the four components of gathering, organizing, summarizing, and reporting as described by Stiggins (2005) and the research questions of this study. The research questions serve as a foundation for this literature review.

1. What are the features of the classroom assessment practices currently implemented to:
    a. gather evidence (assessment data) of student learning in elementary general music in Iowa?
    b. organize student learning data in elementary general music in Iowa?
    c. summarize assessment data in elementary general music in Iowa?
    d. report assessment data in elementary general music in Iowa?
2. What demographics, or teacher characteristics, influence or relate to classrooms assessment practices of gathering, organizing, summarizing, and reporting assessment data in elementary general music in Iowa?
3. What are teacher beliefs or opinions related to assessment and reporting in elementary general music in Iowa?

Included in this chapter are descriptions of both general education and music education classroom assessment practices. The major sections of this review included a general

history of assessment, a description of the assessment practices (gathering, organizing, summarizing, and reporting), and teachers' motivations for assessment as reported in research literature. I am specifically looking for reported data or gaps within the broader scholarly community on a national scale, as well as in Iowa.

History of Assessment

Historical happenings in both general education and music education are the foci of this first section of the review of the literature. This section begins with a brief history of assessment, and then describes literature regarding more recent changes.

History of Assessment in Education

Assessment has played a role in teaching for "centuries" (Guskey, 1996, p. 14) as the ancient Greeks used formative assessment to guide their teaching. Lehman (1968) noted the tests of strength in combat as one of the earliest "performance tests" (p. 4). China was reported as the first documented location that used standardized tests in 2200 BC. For hundreds of years, the Chinese government utilized standardized testing to potential employees interested in joining or reapplying to the Civil Service (DuBois, 1970; Haladyna, Haas & Allison, 1998). Standardized tests are exams that are uniform in every manner (Bracey, 2000).

In the United States, little was known about any grading or reporting prior to the mid 1800s (Edwards & Richey, 1947; Guskey, 1996; Guskey & Bailey, 2001). Either formal assessment was not taking place or occurrences of assessment were not noted at that time. According to Brookhart (2004a), teachers in the United States began using report cards in approximately 1840, modeled after higher education exemplars.

Summative assessment and reporting student progress to parents first occurred in the United States in approximately 1850 (Guskey, 1996). Most schools grouped all students in one room together and few students studied past the elementary grades (Guskey, 1996). As each state passed mandatory attendance laws from the late 1800s to the early to mid 1900s, student numbers increased and students were grouped according to age (Edwards & Richey, 1947). Teachers then began assessing students by making a list of accomplishments attained by each individual student and also noting any remedial work necessary (Edwards & Richey, 1947). Elementary teachers continued the use of such anecdotal records, but early in the 1900s, secondary teachers changed to using numbers and percentages to document students' accuracy and achievement (Kirschenbaum, Napier, & Simon, 1971).

Standardized tests didn't appear in the United States until the middle of the 20th century (Calfee & Masuda, 1997; Johnson & Reed, 2002). Every aspect of the exams was consistent: the directions, the questions, the format, the allotted test time, the testing conditions, the answer sheet, and the scorer training (Bracey, 2000). The goal of standardized tests was producing scores that are interpretable and without bias (Elford, 2002). Standardized tests are either norm-referenced or criterion-referenced. Norm-referenced tests are designed to report results in comparison to a sample of similar students. Criterion-referenced tests report results with reference to a specific set of behaviors, criterion, or standards (Boyle, 1974; Bracey, 2000; Elford, 2002; Gallagher, 1998; Gallavan, 2009b; Labuta, 1974).

Changes in the twentieth century. The twentieth century was a time of great change in educational assessment for many reasons. The Industrial Revolution, increased immigration, and urban expansion were key to the increase of standardized test usage in the United States, as literacy became more important and industries were very interested in well-educated employees (Black, 1998; Perrone, 1977). Standardized exams for elementary and secondary schools were first introduced in Boston in 1845. As the number of immigrants in the United States continued to increase, "the standardized test became a way to ensure that all children were receiving the same standard of education" (Office of Technology Assessment, 1992, p. 1). The very first college entrance exam was utilized in 1851 at Harvard University (Black, 1998).

The use of standardized testing in the schools continued to increase as the student population increased in the early 20th century and school districts nationwide sought objective evaluations of students and schools (Calfee & Masuda, 1997; Tyack, 1974). Standardized testing continued to grow in use within the schools as district and state administrators discovered increased inconsistency and unreliability with individual teachers' marks and ratings of students (Lincoln & Workman, 1936).

The early years of the twentieth century were the beginning of the letter grade vocabulary used today of ranges *A* through *F* (Guskey, 1996). The goal in using these ranges was to provide a larger span for teachers, where an A signified an excellent score range of approximately 90 to 100%, rather than pinpointing an exact score, and thus a fairer distribution of grades with less subjectivity. In the 1930s, popularity of grading

"on the curve" increased where students were "rank-ordered according to some measure of their performance or proficiency" (Guskey, 1996, p. 15).

Questionable practices. Along with the evolution of assessment practices came challenges as well (Starch, 1913; Starch & Elliott, 1912; Wrinkle, 1947). Wrinkle (1947) described the "greatest single change" (p. 50) in grading practices in the first half of the 20th century to be the grouping of percentage grades into ranges, such as 80% to 90%, and the addition of letter symbols for grades, namely A, B, C, D, and F. In 1939, William R. Ross (as reported in Wrinkle, 1947) delivered a national survey to elementary and secondary schools. The survey indicated that 80% of the schools utilized an average of four to seven grading categories, with elementary ranging from 0 to 11 categories, and secondary ranging from 2 to 10 categories. Wrinkle (1947) further described schools in the 1930s that used two-category scales, utilizing either satisfactory/unsatisfactory or pass/fail. The disputability of these categories was the main focus of research by Starch and Elliot in the early part of the 20th century as they questioned the reliability of the grading criteria and factors contributing to the categorical grading practices (Starch, 1913; Starch & Elliott, 1912; 1913).

Objectivity was also a concern related to assessment (Starch & Elliott, 1912). Shortly after the turn of the twentieth century, Starch and Elliott (1912) investigated objectivity of teacher evaluations. Their study found very low reliability in grading practices of high school teachers. A follow-up study by the same researchers in 1913 found similar low reliability in grading practices in different subject areas. Studies such as these influenced many teachers and districts to improve reliability and to change from

percentage scores to a three- or five-point scale with descriptive words used along the continuum, such as *excellent*, *average*, and *failing*.

Impact on learners. As teachers and school districts continued to evolve in grading practices, grading techniques also had an impact on learners. Page (1958) conducted a study where some grades or marks were accompanied by positive teacher comments, while others were not. Researchers found that grades accompanied by additional comments carried a beneficial effect on student learning (Guskey, 1996; Stewart & White, 1976). Parsons (1959) studied elementary teachers' grading and reported a combination of academic and nonacademic criteria in grading and reporting in the late 1950s. Parsons found that test and assignment scores as well as nonacademic criteria, such as behavior, attitude, and effort, were utilized as academic criteria within assessments. Students' behaviors, attitudes, and effort were impacting their academic assessments and academic careers.

Outside influences. Many factors influenced educational assessment. Popham (2001) described the shift in public opinion of education in the middle of the twentieth century toward discontent. Dissatisfaction grew out of many factors with student learning and professional teaching taking the blame. The Russian successful launch of Sputnik in 1957 was a major event that spurred an increased yearning for scientific studies and pursuits in the United States. Further scrutiny arose with continued social discriminatory actions against minorities at this point in history. Reliability and validity questions arose as to whether or not all ethnicities were equally serviced when taking standardized tests. Furthermore, continued problems stemming from the Great Depression continued to

influence the general public's opinion on education. Large numbers of citizens were questioning the relationship between testing outcomes and income and working potential (Popham, 2001).

Considering all of these scenarios, education was thought to be not only a means to greater educational gain, but also supposedly to be a precursor to social gain (Popham, 2001). These numerous factors then impacted and increased the doubts, suspicions and misgivings of the general populations toward education. Popham also reported that teachers became the target of much of the public's educational dissatisfaction. Assessments in schools began to increase in number as many individual states' legislatures established testing to take place to assure minimum competency. As much as the tests measured students' progress, the tests were also intended to measure teacher quality as well, much in reaction to public discontent (Popham, 2001).

The Elementary and Secondary Education Act (ESEA) of 1965 was the first federal funding of money to school districts where steps were being made in support of positive student growth (Johnson, 2002). The determining factor, however, for funding revolved around assessment in providing proof that students were improving (Popham, 2001). Large-scale standardized testing was the focus of research studies throughout the late 1970s and early 1980s (Airasian, Kellaghan, Madaus, & Pedulla, 1977; Kellaghan, Madaus, & Airasian, 1982; Lortie, 1975; Rudman, Kelly, Wanous, Mehrens, Clark, & Porter, 1980; Salmon-Cox, 1981; Sproull & Zubrow, 1981; Stetz & Beck, 1979). Throughout history, as any national crisis has arisen, educational accountability has also increased. The public has repeatedly jumped to the assumption that had educational

excellence increased, perhaps the national crises could have been averted. Popham (2001) pointed out the correlation when the need for accountability increased so did the need for "objective evidence" (p. 107) that quality instruction has led to increased learning. Thus, a simultaneous rise in the use of assessments in education to prove learning has indeed increased.

Accountability in education has continued to grow through the years, but accentuated attention to accountability in education – and the shortcomings of education--have historically occurred (Spring, 1998). When the general public was concerned with the funding of education or when education has experienced a downfall in the United States' international global competition, attention to assessment in education has grown (Simanton, 2000). Hoffer (2008) stated that increased attention on assessment in education has often been related to governmental influences, not necessarily teachers' increased interest in improving student performance.

President George W. Bush signed the No Child Left Behind Act into law in 2002. This was a reauthorization of the Elementary and Secondary Education Act but more all-encompassing, with more emphasis on the importance of assessment of student learning. All students under the law were to be tested, and the test results were connected to funding and other supports. This was monumental emphasis on testing for kindergarten through twelfth grade students (Abeles, 2010; Branscome, 2005)

Recent studies. More recent studies have focused on describing the types of classroom assessment methods utilized in the United States (Adams & Hsu, 1998; Gullickson, 1985; McMillan, 2001; Stiggins & Bridgeford, 1985). Gullickson (1985)

conducted research where 295 teachers in a rural, Midwestern state were surveyed on their student assessment practices. Gullickson was particularly interested in whether these practices varied by grade level and/or subject area. Elementary teachers were found to use a variety of assessment techniques, primarily informal and observational. Secondary teachers were found to rely on tests. Elementary teachers were found to use more commercially prepared tests than secondary teachers who indicated more uses of original tests.

Stiggins and Bridgeford (1985) conducted research surveying 228 teachers in eight different school districts throughout the United States. These researchers, through the use of a survey, sought both an inventory of assessment practices and also the perceptions and ideas linked to the practices. This data showed that the majority of teachers used original documents for assessments and indicated a need for assistance in improving or changing assessment efforts. In both studies, Stiggins and Bridgeford (1985) and Gullickson (1985) found that increased concern about assessment escalated as the age of the students increased.

A study in 1998 by Adams and Hsu found that elementary teachers primarily used observation in the classroom. Having surveyed 269 elementary teachers in Florida, they found that teachers' attitudes, beliefs, and concepts influenced any changes being made in curriculum or instruction in mathematics. Several further studies (Bateson, 1990; Griswold & Griswold, 1992; Gullickson, 1985; Nava & Loyd, 1992; Wilson, 1990) showed a grade-level effect on grading practices. Elementary teachers utilized, on

average, more informal assessment practices and most often used observation as compared to teachers at the secondary level who utilized more formal methods.

McMillan (2001) surveyed 1,483 secondary teachers in Virginia to seek information regarding validity and the basis for grading practices, specifically in regard to what factors are considered when figuring grades. Similar to Parsons' (1959) study, he found that teachers used a variety of nonacademic factors when grading, such as efforts, behavior, and attitude. Many further research studies have found great variation among assessment practices (Brookhart, 1993; 1994; Frary, Cross & Weber, 1993; Nava & Loyd, 1992; Pilcher-Carlton & Oosterhof, 1993). Numerous authors, however, found a significant difference between what teachers consider "best practice" (Stiggins, 1998, p. 4) and what was actually occurring (Barnes, 1985; Cizek, Fitzgerald, & Rachor, 1995; Cizek, 1996; Frary et al., 1993; Haladyna, 1999; Manke & Loyd, 1990; Stiggins, 1998). To summarize the research on practices in the 1980s and 1990s, Guskey (1996) stated that teachers continued to strive for excellence in educational assessment, but commonly missed the mark due to continued fluctuation or inconsistencies.

Brookhart (2004a) published an extensive review of literature relating to classroom assessment with the aim of finding the relevant ideas associated with teachers' practices. She found three prevailing functions: instruction, management, and assessment. Brookhart found that the overlapping of the three functions has provided a rich foundation for understanding.

There have been numerous studies in assessment over the last 50 years (Guskey, 1996). Guskey identified many studies where researchers have identified many different

practices and purposes in assessment and grading, but no one method or practice that could achieve all of the purposes well (1996).

History of Assessment in Music Education

Music educators, not unlike the general classroom teachers, have sought to evaluate students' achievement in music education. Students in England were assessed on solfege singing in the 19th century (Osborne, 1983). In America, during the early 1920s, following the beginning of organized bands in high schools, high school bands could compete in instrumental performance competitions that were coordinated by the Music Educators' National Conference (MENC, now called the National Association for Music Educators). Many band directors viewed the qualitative critiques from judging at these contests as substantial assessment practices and performed no other musical assessment. Some band directors would use the overall contest rating received by the entire ensemble to each individual band member as their grade (Lehman, 1992; McFarlin, 1965; Schleuter, 1984; Simanton, 2000).

Zaymeyer (1959) described post-World War II music grading practices as ensemble directors' views on participants' attendance, participation, and rehearsal skills. The practice described students beginning each rehearsal with a certain number of points and losing points for each infraction or improper choice. Another practice in the early 1970s was documenting weekly practice habits of secondary ensemble members. Each student was graded on "practice, tone, technique, rhythm, articulation, intonation, phrasing, breath control, embouchure, and band music" (Senty, 1971, p. 22). The students' progress was then communicated to the students and their parents. A similar

method was described by Pace (1972). Students were expected to complete a checklist of expectations and tasks in order to pass the music class each semester.

Lack of credibility. Music educators were lacking clear objectives, a method for measuring student progress, and a credible way to evaluate students' overall work in music (Ellis, 1963). Grades, if given, were assigned based on varying subjective attributes. Ross (1975) described various grading scales where the average letter grade of C was given to all music students, unless additional assignments were completed.

The 1980s was also a decade that saw many music educators rushing to assess student work, seeking credibility and a new means to meet the increased demands on music and maintaining music as a quality component of the curriculum (Walker, 1998). Greer was noted as one of the first music educators to consider music education using behaviorist principles (Duncan, 2009). In 1980, Greer strongly suggested communicating students' progress in music to families. The drill and practice philosophy was part of Greer's idea of strictly defining the musical objective, teach the objective, provide feedback on the objective, and then change the objective if needed. This process was simply suggested to be repeated until the students achieved mastery of the objective.

Zdzinski (1996) described numerous formative and summative assessment methods that have been developed for secondary large ensemble rehearsals in the 1980s and 1990s. Further secondary music student assessments have included audio and videotapes and also rating scales (Carlin, 1996; Cope, 1996; Killian, 1995; MENC, 1996b; Robinson, 1995; Swanwick, 1994). Studies found that high school music teachers, however, were basing grading practices primarily on attendance (Lehman,

1992; McClung, 1997). Regardless of the age level or medium, Shuler (1996a) summarized that music teachers have, throughout time, continually been "insecure about assessment" (p. 10). Furthermore, music educators have continually argued the high level of personal expressiveness in students' performance of music, which leads to much difficulty in objectively evaluating (Radocy, 1986; Wesolowski, 2012).

Standardized tests in music education. Many music educators attributed the lack of respect by other content specialists to no utilization of standardized testing in music (Shuler, 1996a). In seeking credibility, many teachers and districts began implementing some of the more popular standardized tests in music, such as the Music Achievement Test by Colwell (1969), the Musical Aptitude Profile by Gordon (1965), or the Iowa Test of Music Literacy, also by Gordon (1970). Unfortunately, by labeling these types of tools as assessments, many teachers avoided these tests. Because of teachers' negative attitude toward assessment and testing, they minimal interest in assessment (Shuler, 1996a).

Although not well received by all music educators, numerous standardized music tests, focusing on aptitude and achievement, were published in the last fifty years (Kotora, 2001). Most were developed and published 40 to 50 years ago, and some have been updated more recently (Leonhard & House, 1972; Simanton, 2000). The tests were intended to both identify students with extraordinary promise in music to encourage enhanced musical activities, and also to benefit all students by the diagnosis of strengths and weaknesses to adapt instruction to meet students' needs (Gordon, 2007).

The aptitude tests were developed to measure students' capacity or potential for achievement (Lehman, 1968). Aptitude generally encompasses one's natural genetic abilities and any maturation or significant environmental influences that are not from formal instruction (Lehman, 1968; Leonhard & House, 1972; Radocy & Boyle, 1988). Carl E. Seashore, a prominent music psychologist, developed the Seashore Measures of Musical Talents (Seashore, Lewis, & Saetveit, 1960) with the oldest version first used in 1919 as the "earliest and most recognized" (p. 9) music test (Kotora, 2001). The Seashore tests were out of print by 1994 (Boyle & Radocy, 1987). Additional aptitude tests on music through the 20th century included Drake Musical Aptitude Tests (Drake, 1957), Wing Standardized Tests of Musical Intelligence (Wing, 1961), the Musical Aptitude Profile (Gordon, 1965), the Measures of Musical Abilities (Bentley, 1966), the Primary Measures of Musical Audiation (Gordon, 1979), and the Intermediate Measures of Music Audiation (Gordon, 1982). Lehman (1968) described how interest in testing of musical aptitude increased in the late 1920s and began to decrease after the 1960s.

Standardized tests of musical achievement have also existed throughout the years (Harrison, 1983). Boyle and Radocy (1987) defined music achievement as the "general musical knowledge, knowledge of notation, aural-visual skills, aural skills, and compositions as well as performance" (Boyle & Radocy, 1987, p. 157). These authors described that previous academic achievements may determine what a person can "do in the present" (p. 157). The many music-standardized tests on achievement are Watkins/Farnum Performance Scale, (Watkins & Farnum, 1954; 1962), Farnum Music Tests, (Farnum, 1969), Music Achievement Tests (Colwell, 1969; 1970), Indiana-Oregon

Music Discrimination Tests (Long, 1970), Simons Measurements of Music Listening Skills (Simons, 1974), and the Silver Burdett Music Competency Tests (Colwell, 1979).

The two most comprehensive, utilized, and available music tests for elementary were the Music Achievement Tests (Colwell, 1969; 1970) and the Iowa Tests of Music Literacy (Gordon, 1970), although neither test directly measured performance (Simanton, 2000). Richard Colwell developed the Music Achievement Tests in four components. The first two components were published in 1969 and the last two components in 1970. Geared for students in grades 4 through 12, the Music Achievement Tests (M.A.T.) were a plethora of listening exercises from which teachers would choose the most appropriate exercises with relation to their own teaching and curriculum. These tests were also norm-referenced and based on data norms from 1969. These achievement tests were significant because they were seen by educators as a valid attempt to objectively assess factual knowledge accumulated by students from instruction and experiences (Lehman, 1968).

In addition to specific aptitude and achievement tests, other tests attempted to measure both aptitude and achievement (Harrison, 1983). The Iowa Tests of Music Literacy (Gordon, 1970) were first published in 1970 and designed for students in grades four through 12. Known as the ITML, these tests measured music aptitude using both melodic and rhythm aural identification and music achievement using music terminology and reading (Gordon, 1970, 1991). A revised and updated edition of the test was published in 1991 and relates to Edwin Gordon's *Jump Right In* (2000) elementary music curriculum. This norm-referenced exam was less difficult and time consuming to score,

but Radocy (1989) described the updated test as still using the original 1970 national sample data for normative comparison.

Law (2012), to the contrary, described many limitations to the standardized tests in music and rationale for their decreased or limited use. Many of the tests were designed specifically to measure music aptitude, abilities, or achievement prior to music education. So using them in a classroom as a summative tool would not be suitable. Additionally, many of the standardized music tests were created for young children, thus not appropriate for older children. Another finding by Law was the rhythmic and timbre inconsistencies of live performers providing the musical material required for students to complete the exams. The recordings provided for the exams are dated and very low quality recordings. Furthermore, test design as well as reliability and validity have proven problematic through time.

National Standards in Music Education. Assessment was slowly becoming a part of music education publications beginning in the 1970s (Brophy, 2000). Although not a priority, assessment was referenced along with curriculum goals and guidelines on the national level (MENC, 1974). A change in priorities occurred in the early 1980s when the National Commission on Excellence in Education (1983) published a report in 1983 entitled *A Nation at Risk: The Imperative for Educational Reform* (Education Commission of the States, 1983), which explained the necessity for every academic subject area to delineate standards with measurable outcomes (Duncan, 2009; Giordano, 2007). In the years following, several streams of advocates worked independently and finally joined together in 1990. By 1991, MENC published a report from multiple efforts

entitled, *Growing Up Complete: The Imperative for Music Education.* This was an important factor in a joint effort to have the arts included in the *Goals 2000: Educate America Act* (Mark, 1995). At the same time, the National Council on Education Standards and Testing (1992) called for voluntary standards from all subjects, including music. Through MENC (1994a), a task force of professional educators in all arts (music, theater, dance, and visual arts) worked together to develop the National Standards for the Arts Education, published on March 11, 1994. By the end of March, then President Bill Clinton signed them into law (P.L. 103-227) as a portion of the *Goals 2000: Educate America Act*. The arts were identified as a core subject among all of the educational disciplines for the first time in American history (Abrahams, 2000; Byo, 1997; Byo, 1999; Kotora, 2001; MENC, 1996a; Riley, 2009; Schwartz, 1996). The law itself was initiated from the six National Educational Goals established by President George Bush and the 50 state governors in 1990, as they were focusing on accountability in education (Kotora, 2001; Simanton, 2000).

The National Standards and Title II of *Goals 2000* defined what every child should know and be able to do to demonstrate competence in music (Lehman, 1996). The National Standards for Music Education state that all students in elementary, junior high, and high school should partake in various musical activities throughout their educational career (Lehman, 1993). The National Standards for Music Education included singing, playing instruments, improvising, composing, reading or notating, listening or analyzing, evaluating, understanding relationships between music and other disciplines, and understanding music in relation to history and culture. The Standards were divided into

graded segments of Kindergarten to fourth grade, fifth through eighth grades, and ninth through twelfth grades (Lavender, 2000). The Standards identified two areas of standards: content and achievement. The achievement standards also established the "understandings and levels of achievement" (p. 80) expected of students in each competency (Schwartz, 1996).

The Standards, although not mandatory, were intended to provide a model for all states to use as a basis for developing their own standards and to provide a foundation for all music curricula, as well as a means for assessing student achievement and performance (Brophy, 2003; Consortium of National Arts Education Associations, 1994, Kotora, 2001; Lehman, 2000; Roeber, 1995; Schmid, 1996, Shuler, 1996b). Within the first two years of implementation, 36 individual states began devising state level standards in the arts or revising existing state standards, using the National Standards as guidelines (Lehman, 1996). Ambach (1996) found data from surveys by the Council of Chief State School Officers (CCSSO) and others indicated use of the standards in 40 to 43 of the states as "substantial" (p. 7) with much variation in use, adoption, and application.

The Standards were and are a means by which music educators can enhance their music teaching through increased accountability and assessment (Welter, 1993). With regard to assessment, students in grades Kindergarten through fourth and grades five through eight were expected to meet specific age appropriate achievement standards. Students completing one or two years of secondary music were expected to achieve at the proficient level. Students enrolled in music classes for three or four years in high school

were expected to meet the advanced level standards. Shuler (1996b) concluded that the National Standards had a "profound" (p. 87) impact on student assessment, as many major arts initiatives were outgrowths of the National Standards, and music teachers finally had a foundation on which to base evaluations.

With these expectations, music teachers soon looked to the national level for guidance on implementation and assessment in the classroom (Lehman, 2000; Salvador, 2011). The Music Educators National Conference made efforts to provide guidelines for teachers attempting to implement the Standards into curriculum (MENC, 1994b). Several publications of such by MENC were published in 1994, which included editing committees or chapter submissions from leading music education experts from across the country.

By 1996, the MENC (1996b) published a basis of guidelines for assessment in music, which stated that assessment should be standards-based, reflective of music skills and knowledge, supportive of learning, reliable, valid, authentic, and open to review (MENC, 1996b, p. 7-9). The National Standards did not, however, include any specific steps for assessment or grading (MENC, 1996b).

> "That is left to the states, local districts, and individual teachers. Because assessment procedures must be based on instructional procedures, differences in assessment procedures are expected, as well as differences in methodology. Teachers should feel free to devise alternative assessment procedures that will work in their situations" (p. 12-13).

Lehman (2000) further stated that the implementation of the standards in 1994 was an opportunity for increasing the expectations associated with assessment and that assessment should be coexistent with the National Standards in teaching.

By 2003, MENC published a book and coordinating CD entitled, *Benchmarks in Action: A Guide to Standards-Based Assessment in Music*, edited by Carolynn Lindeman. This publication was designed to aid teachers in "assessing students' progress in music in grades K-12" (Lindeman, 2003, p. ix). The book provides basic, proficient, and advanced levels of achievement on all nine standards for all grade levels. Also included was a CD recording of the different levels as well as excerpts for the listening examples (Lindeman, 2003). Although this publication was aimed at assisting music educators with the daunting task of assessing the students' performance and attainment of the standards, it requires a large working knowledge of the standards in order to be utilized. There are many overviews and guidelines included, but the implementation of the ideas from this book would require both study time by the teachers and class time for implementation in instruction for the students (Lindeman, 2003).

Teachers' implementation of the National Standards in Music Education. The publication of the National Standards in Music Education created a potential for much change in music education, dependent on how teachers of music received and implemented the information (Colwell, 2008). Many studies have been conducted in search of music teachers' practices for implementing the National Standards in Music Education in the schools.

Since 1994, studies have focused on abilities of teachers to implement the National Standards (Byo, 1997), attitudes, perceptions and abilities of pre-service music teachers to implement the National Standards (Abrahams, 2000; Riley, 2009), amounts of class time spent with regard to the National Standards (Orman, 2002), attitudes of music

teachers regarding the implementation of the National Standards (Louk, 2002), and graduate class content on the National Standards (Bell, 2003).

Byo found that music teachers (N=122) were able to implement the standards with ease and with little assistance and that there might be a connection between teacher training and abilities to implement the National Standards for music education. The study also revealed that some of the standards were implemented with more ease and less effort than others, namely singing, analyzing music, and history and culture. Abrahams (2000), found both institutions in his study had made several changes to implement the standards. Abrahams found that pre-service teachers did not know about the National Standards until college and that their abilities to musically improvise were low. The students in this study believed that their knowledge of the standards would be beneficial in future career searches.

Orman's (2002) study revealed that the nine standards were taught or implemented unequally in music classes and that data indicated the standards that received the most content time in classes were singing, playing instruments, and reading and notating. The standards that were allotted less time were those that required more creative and artistic skills from the students, namely evaluating, composing, and improvising. Louk (2002) similarly found the most important standards, in the opinions of the respondents, to be reading and notating music, history and culture, and performing on instruments. The least important standards, as indicated by the data, were reported as evaluating, improvising, and composing. Furthermore, Louk's study revealed significant

differences in the survey responses, where the teachers documented how much time and effort they put into the implementation, and the actual observations of time in the classes.

Bell (2003) found that the teachers were aware and had a working knowledge of the National Standards, however, the different school districts displayed inconsistent applications and support systems. The standards reported most difficult to implement were the singing standard (36%) and the improvising standard (36%).

Most recently, the results of Riley's (2009) study found that 100% of the music education students surveyed agreed that they were aware of the standards after the course, where only 27% had indicated so prior to the course. The students were most comfortable with the reading and notating standard, and the evaluating standard. The students were least comfortable with the composing standard, the understanding relationships standard, and the improvising standard. These studies are significant as the data indicated continued lack of clarity with regard to the National Standards in Music Education.

Practices in Gathering Assessment Data

This research study was seeking data with regard to how elementary music teachers in Iowa gather, organize, summarize and report assessments, and specifically how music teachers have been and are currently conducting these tasks. Seeking a thorough description of how teachers gather assessment data, this next section will describe the types of assessment data gathering practices used by teachers as reported in the literature.

Assessment was defined as a "process" (p. 2) beyond just one tool or measurement (Farrell, 1997). Duerksen (1995) reported that music teachers utilize a variety of assessment tools in the classroom. Several authors noted the importance of knowing about, or learning about, a multitude of assessment strategies (Lehman, 1997; Schultz, 2002; Welter, 1993). Guskey (1996) stated how teachers have options of utilizing different assessment techniques "individually or in combination" (p. 84), always striving to obtain the most critical information about each student's learning.

The literature also documented the importance of using the appropriate tools in the classroom for measuring student achievement and growth and keeping accurate records (Johnson, Dupuis, Musial, Hall, & Gollnick, 2002; Lehman, 1998; Mierzwik, 2005). Not only must teachers consider what type of assessment design to use, teachers also must choose between an existing assessment format or creating an assessment of their own. Many researchers have documented that 90 to 95% of teachers in general have, through the years, used tools they have created themselves (Dorr-Bremme, 1983; Gullickson, 1982; Newman, 1981; Stiggins & Bridgeford, 1985; Yeh, 1978).

Duncan (2009) noted the importance of using a "variety of assessment methods to gain a complete picture" (Duncan, 2009, p. 14) of student music learning and cautioned music educators not to limit assessment practices to one type. Similarly, the Council of Chief State School Officers (CCSSO, 1997) stated that teachers should include a variety of assessment tools and formats, to assure that assessment efforts are a true reflection of learning--both quantification of scores of achievement, as well as students' critical judgment and application of their knowledge.

Specific Types of Data Gathering Tools

Teachers in both general classrooms and music classrooms utilize many different assessment tools or formats as a means of gathering data on student progress. The variety of instruments specifically described in the next section of this literature review includes examinations, projects, assignments and homework, observations, performances, portfolios, and rubrics as described in the literature. Stiggins (2005) emphasized the importance of choosing the appropriate assessment methods that matches the expectations established in the objectives. McTighe and Ferrara (1998) recommended letting the choice of assessment format to be determined by the (a) overall purpose of the assessment; and (b) by the age of the student. Teachers have numerous choices to consider when assessing students.

Examinations. The majority of general classroom assessment tools are tests or quizzes (Slavin, 2003). These tests and quizzes can include a myriad of tools, as they may be original or they may be produced and published by other sources (Chittenden, 1991). Mierzwik (2005) described the importance of tests and quizzes, as the goal of these tools continue to be to show how well students have mastered the curriculum.

Selected response, a common type of exam, began use in the 1920s. In selected response exams, students were and are asked questions with a range of responses with the task of selecting the correct or best answer from the options given (Stiggins, 2005). Examples of selected response assessments include multiple-choice questions, true-false questions, and matching. Selected response assessments are easy to score, objective,

provide the opportunity to cover much information with limited time and space, and can measure knowledge, comprehensions, and application (Taylor, 2003).

Projects. A project as an exhibition that illustrates more than a final product but rather the many steps required in achieving the final product. Many different answers or strategies can be acceptable. Completion of a project may require much persistence and time. Projects may demonstrate more complex applications of students' skills (Farrell, 1997).

Assignments and homework. Mierzwik (2005) noted many benefits of assigning homework to students as practicing skills may add to mastery. While students are practicing, teachers are also monitoring learning (Johnson et al., 2002). Management, however, of the students' work and scores can be overwhelming to teachers. Parents, alike, can view homework assignments as too much work or even too confusing (Mierzwik, 2005). Mierzwik (2005) described a "homework policy" (p. 24) where teachers clarify expectations and make the work manageable for the students and parents. Aligning the homework policy with district standards and clearly documenting the policy were also recommended.

Observations. Observations are objective notes that teachers make based on their visual perceptions of student work seeking to gather data to understand a learning situation or behavior (Chittenden, 1991; Taylor, 2003). When using observational techniques, teachers must look for "cues" (Chittenden, 1991, p. 25) that signal students' learning through their "interests, thinking, and relationships" (p. 25).

Observations were further described as the “favored teacher technique” (Salmon-Cox, 1981, p. 632). Many studies have noted the high frequency of observation techniques in the classroom (Herman & Dorr-Bremme, 1982; Kellaghan et al., 1982; Salmon-Cox, 1981; Stiggins & Bridgeford, 1985). Taylor (2003) and Abeles (2010) agreed how data gathered from observations may be used to make decisions with regard to the future of individual students, classes of students, and teaching methods or instruction.

Mierzwik (2005) explained the importance of documenting observations with clear organization. All observations should include each student’s name with vertical columns for each task observed. At the top of each column should be the title of the observed task and either a number of possible points or an explanation of other marks entered. Rating scales or checklists can quickly organize the observational data.

Rating scales are means to quantify data collected from observations (Taylor, 2003). Rating scales list specific attributes of student performances and designate a total number of points possible for each attribute (Simanton, 2000). Following the performance or observation, rating scales can assist in collecting data of students, with the appropriate format or organization.

Checklists are another technique for organizing observational data (Harrison, 1983). A checklist has been and is used to document a seen or unseen behavior through observation, looking for specific “dimensions, elements, activities, characteristics or behaviors” (Farrell, 1997, p. 15). It can be used to observe an individual or a group of students. It is literally a list of objectives that can be plainly checked off when seen or

observed. Kriske and DeLelles (2005) suggested using a seating chart in the music classroom as a type of checklist where each seat indication on paper also includes several small boxes for checks, plusses, and minuses. Taylor (2003) suggested categories for tallying observances of certain behaviors or skills, such as on task behavior, singing in tune, or critical observations. Teachers may observe how many times an event occurs, the duration of certain events, or how often certain behaviors occur.

Performances. The literature referred to a judged performance assessment as an authentic assessment or alternative assessment (Berman, 2008; Farrell, 1997; Foley, 2001; Shuler, 1996b; Slavin, 2003). The aim of performance or authentic assessments in either the general classroom or the music classroom continues to be providing real world tasks that require knowledge and skill to perform. Some performances may be for a large audience with a known purpose, location, and time. Series of performances can illustrate growth or maturity following much practice or rehearsals (Farrell, 1997).

Farrell (1997) emphasized the importance of process as well as the final product, and thus the importance of considering the performance involved. Also visible in a performance are the many steps that students take leading up to a final performance or product. Observing, journaling, or asking questions about such steps can also indicate achievement and growth. The overlap, here, with both projects and observations were evident.

Performance assessments are based on the teacher's observation of the performance or any products created during the performance. Students are expected to actually "carry out" (p. 66) the assigned behavior. Performance assessments were

described as another way of gathering evidence of learning (Stiggins, 2005). With authentic assessments, students are actively participating in the assessment process, for example musical performances or compositions (Boyle, 1996).

In music, authentic assessment has been described as a practical technique of evaluation, since music is a performing art (Kotora, 2001; Hanzlik, 2001; Hill, 1999; McClung, 2000). The focus of authentic assessments was described as focusing more on behaviors generated by the student (performance) rather than choosing a response (Lehman, 1996). Authentic tasks in music include the following: performances, compositions, investigations, analyses, demonstrations, written or oral responses, journals, and portfolios (Lehman, 1996, p. 111).

Cope (1996) illustrated the practicality of such ideas with the close relationship of a performing art and measurement techniques that focus on more than just answering questions about music. The focus was and is the collection of performance recordings and projects into a portfolio format that may indeed provide a clear picture of achievement growth (Farrell, 1997). Nierman (1996) pointed out how the active participation of the students required in performance assessments requires them to apply their knowledge in creative settings, and not just regurgitate memorized content.

Portfolios. A portfolio was described as a "purposeful collection of student work assembled to provide a representation of student achievement" (Stiggins, 2005, p. 320) with the intent to "tell the story" (p. 320) of students' learning and progress. The contents of a portfolio has been considered "artifacts" (Hill, 2008, p. 61) of documentation (Hill, 2008). Several authors noted the variety of works that can be

included in a portfolio (Cizek, 1997; Farrell, 1997; Goolsby, 1995; Robinson, 1995). Portfolios can be a culmination of one subject area or class, or a combination of work in many classes or even grade levels. It is important, though, to note the difference between portfolios of student work and a simple file of work by students. The portfolio was and is intended to contain purposefully selected items by the student that demonstrate growth and learning, not just a collection of all work completed (Mierzwik, 2005).

The goals of a portfolio are to illustrate student growth through time, with consideration for both breadth and depth of student achievement, all with the ultimate goal of showcasing either progress of the student or simply the best work of the student (Kelly, 2001). Portfolios can demonstrate higher order thinking (Kelly, 2001). Mierzwik (2005) also suggested student reflections of work to be included in portfolios. When teachers review students' work within a portfolio, they are not just looking for mastery, but also growth (Goolsby, 1995). The variety of works can contribute to a clearer overall illustration of students' growth and development over time (Kotora, 2005; TenBrink & Cooper, 2003)

The music teacher may include a scoring guide, or rubric, with a portfolio of student work to evaluate each work contained in the portfolio (Nitko, 1996; Simanton, 2000; Stiggins, 2005). A rubric was defined as a scoring tool used by teachers to outline varying degrees of adequacy within any given assignment (Hickey, 1999). Each level of proficiency must include a thorough description of performance for that level. Each descriptor contains brief statements or sentences that define the particular level. Rubrics establish consistency and reliability in the evaluation of portfolios in reducing

opportunities for subjective judgments (Fransen, 1998; Goolsby, 1995; Hickey, 1999; Kelly, 2001; Robinson, 1995)

Goolsby (1995) outlined the evolution of portfolios in music education and attributed the initial "support for alternative assessment from policymakers" (p. 39) as coming to be because of the gradual realization that music may not be properly or accurately assessed in typical, namely standardized, means. Fransen (1998) furthermore reported that portfolio use in music education was beneficial and potentially useful in the music classroom. She investigated 20 students during a music curriculum unit and observed and interviewed the students as they assembled and showcased their music portfolios. The students in this study enthusiastically completed all of the tasks involved in the research study and articulated what they had learned during the unit of study.

Rubrics. Many of the above-mentioned tools require scoring guidelines, or a rubric, to evaluate students' work (Carr & Harris, 2001). Rubrics are an organized list of the required criterion for the work itself. Not unlike a rating scale or checklist, a rubric is typically divided into categories. A rubric, however, adds specific expectations for varying levels of performance or achievement. Students can refer to the rubric during preparatory steps and teachers can use the rubric for scoring procedures (Whitcomb, 1999). Rubrics with detailed rating scales can assist music educators in moving beyond simply collecting performance samples but also assessing the works with as much objectivity as possible (Nierman, 1997).

Several authors have devised rating scales specifically for performance rubrics in music at the secondary choral or instrumental level and presented them in the literature

(Cope, 1996; Killian, 1995; Matheny, 1994; McPherson, 1993; Robinson, 1995; Russo, 1988; Swanwick, 1994). Examples of elementary music rubrics are included in *Performance Standards for Music: Strategies and Benchmarks for Assessing Progress Toward the National Standards* (MENC, 1996b). The exemplars include detailed definitions of each performance criterion at various levels of proficiency with detailed descriptors. These examples are based on the National Standards and include further strategies for implementation.

Frequencies of Gathering Practices in Music

With so many diverse options for gathering assessment data, a logical query would be how often music teachers are utilizing each tool in general music classes (Talley, 2005). One particular research study investigated the frequency of assessment in elementary general music classrooms in Michigan (Talley, 2005). This lone recent study surveyed elementary music teachers regarding how often assessments are administered at each grade level, and specifically what type of tool was used. The data in this descriptive study revealed a reluctant and hesitant response, with a majority of the respondents indicating they were not required to assess. Talley found that many of the respondents did not frequently assess their students and some indicated no assessment occurred. The study also revealed an increase in the number of assessments given each year, as the students age increased, with the youngest grade (Kindergarten) revealing the lowest number of assessments delivered. Rating scales and rubrics were reported as the most often used assessment tool with 46% frequency.

A study by Barkley (2006) also investigated the practices of elementary general music teachers' strategies, frequencies, influences and attitudes with and about assessment. Observation was the most common assessment technique indicated in the descriptive research study survey of 255 elementary music teachers in Michigan. Portfolios were reported least frequently used. A similar study was conducted by Hepworth-Osiowy (2004) who surveyed 190 elementary music teachers in Canada. Respondent data indicated a large variety of assessment strategies and practices.

A less recent qualitative study of 3 elementary general music teachers (Nightingale-Abell, 1993) indicated elementary general music teachers used "a variety of teaching techniques" (p. 41). The most common theme found in this study was teacher-made tests, with observations also noted as prevalent.

Practices in Organizing Assessment Data

The second aspect of assessment as described by Stiggins (2005) was organizing assessment data. Regardless of the organizational technique, storage of records and keeping extra copies of data was noted as an important step in the organizational process (Mierzwik, 2005). Organizing the data that has been gathered in both general and music classrooms may be found to be both overwhelming and problematic. Numerous authors, detailed below, have reported that often music educators, specifically, have found challenges in their assessment practices. Other experts have published recommendations for overall assessment organization, as described below. The problems, namely number of students, instructional time, limited training, measurement, and preparation time, and

the suggested solutions, specifically suggestions in efficiency, organization, and saving time, are found in this section of this literature review.

Problematic Organizational Issues

Several sources have provided insight to situations that have hampered organizational practices in music education assessment over time and specifically reported a range of problematic situations in assessment data gathering in elementary music due to large numbers of students, limited instructional time, limited training in assessment techniques, difficulty with measurement issues, and inadequate preparation time. Still other experts have published recommendations for overall assessment improvements. Both the problems and the suggested solutions are found in this next section of the literature review.

Large number of students. Studies showed that many teachers have large numbers of overall students in music classes. It can be very difficult to gather assessment data when considering hundreds of names, faces, and activities to work with and evaluate (Barkley, 2006; Bouton, 2001; Boyle & Radocy, 1987; Brophy, 2000; Brummett & Haywood, 1997; Hanzlik, 2001; Harrison, 1983; Kancianic, 2006; Kotora, 2005; Lavender, 2000; Lehman, 1998; McClung, 1996; McCoy, 1991; MENC, 2007; Nierman, 1997; Nightingale-Abell, 1993; Nutter, 1999; Russell & Austin, 2010; Salvador, 2011; Simanton, 2000; Shih, 1997; Shuler, 1996b; Tracy, 2002). Elementary general music teachers have hundreds of students, typically in many different grade levels. Teachers might not be able to put a face with every name on the class lists, due to so many different students (Harrison, 1983; Lavender, 2000). Occasionally, music teachers may

only have data collected on the students who perform very well or very poorly (Harrison, 1983). Nutter (1999) pointed out how many music teachers serve students in more than one building within a district or town. Barkley (2006) described the idea of "keeping records and grading paperwork" for hundreds of music students to be "overwhelming" (p. 2). Moreover, with so many students to assess at once, the management of students in music classes was also reported as problematic for teachers when assessing individual students (Kotora, 2001; Shih, 1997).

Instructional time. All teachers in all subject areas would have the opportunity to assess more if more contact time was available with students (Rosenshine, 1981). In music classes specifically, the problem of minimal contact time was expressed in the literature (Barkley, 2006; Kotora, 2005; McClung, 1996; MENC, 2007; Nightingale-Abell, 1993; Shih, 1997; Shuler, 1996b, Simanton, 2000; Tracy, 2002). With large numbers of students in an entire elementary building, elementary general music teachers have limited contact time with each individual student as each class must receive music instruction during each week or rotational period (Brummett & Haywood, 1997). When music teachers did not see students often or for long periods of time, delivering quality assessments was reported as a major issue (Barkley, 2006; Kotora, 2005; McClung, 1996; MENC, 2007; Nightingale-Abell, 1993; Shih, 1997; Shuler, 1996b, Simanton, 2000; Tracy, 2002). Scheduling can play a part in the minimizing of instructional time as well. Barkley (2006) documented how some teachers have teaching assignments in more than one elementary building, thus time that could be spent with students was spent traveling between buildings. Teachers assessing students and reporting progress to parents

requires time, both in the classroom delivering the assessments and in preparation of the data to communicate to parents (Barkley, 2006; MENC 1996a). Many teachers see testing as taking up valuable time that was already in high demand (Hamann, 2001; Pontious, 2001).

Limited training. Specific studies in the literature showed a major lack of training in assessment techniques as problematic for music teachers (Boothroyd & McMorris, 1992; Cope, 1996; Kotora, 2005; McClung, 1996; Nightingale-Abell, 1993; Shuler, 1996b). General classroom teachers were also found to be feeling lacking in background or training within assessment, as several studies in education cited a low level of knowledge, training, and interest in grading (Allen & Lambating, 2001; Barnes, 1985; Brookhart, 1993; Carter, 1984; Cizek, 1996; Fautley, 2010; Lomax, 1996; Stiggins, 1988; Stiggins, 1999; Stiggins, Frisbie, & Griswold, 1989; Schafer & Lissitz, 1987; Shuler, 1996b; Stiggins & Conklin, 1992; Thorndike, 1997). In 1991, Wise, Lukin, and Ross's study found that a majority of states did not mandate a course in testing or measurement in order to receive a teaching certificate for any subject. Furthermore, approximately half of their 397 respondents nationwide were reported to believe that any training that they did receive was either partially or totally inadequate (Wise, Lukin, & Ross, 1991). Although this study was in the early 1990s, all teachers--both music and general classroom--who currently are in classrooms and who would have received their original teaching license before 1991, would fall into this category.

Rosen (1999) conducted a multifaceted study of both qualitative observations of nine classrooms and a survey to 30 randomly sampled teachers. This research study

examined the knowledge, perspectives, and practices of teachers in teaching, learning, and assessment. She summarized that teachers' knowledge and use of assessments needed improvement and that professional development should most definitely occur. Abeles (2010) agreed that professional development opportunities in assessment aids music educators' progress in assessment duties. Furthermore, a study by Allen and Lambating (2001) documented the minimal training that actually takes place at the undergraduate level in pre-service teaching. Four-year institutions with education degrees were randomly selected, and less than half (47%) of the schools required a course in measurement. Only approximately one third of the randomly sampled schools (32%) required a general course on assessment; only four percent required a course focused on "informal assessment" (Allen & Lambating, 2001, p. 26); only eleven percent required a course focused on assessment of students with "special needs." Guskey (1996) used the term "few" (p. 2) when describing the quantity of teachers with formal training in the areas of grading or reporting. Guskey (1996) continued to expand on this problematic situation by describing further the lack of direction for teachers who were not happy with their current assessment practices.

With specific regard to music, Kotora (2001) conducted a study of Ohio high school choral music teachers and college choral methods teachers. This study revealed that 66% of the respondents who did have assessment courses did not believe these classes prepared them thoroughly for the classroom. Furthermore, 53% of respondents with graduate degrees stated the same disenchantment with masters' level courses in assessment.

Regardless of training, the literature suggested teachers commonly utilize the types of grading procedures that were previously used on them when they were students (Allen & Lambating, 2001; Britzman, 1991; Ginsburg & Clift, 1990; Holt-Reynolds, 1992; Pajares, 1992). The teachers' "personal experiences" (Allen & Lambating, 2001, p. 8) that they had as a student themselves overrode any training or principles in academic achievement (Allen & Lambating, 2001). After numerous years as a student, studies showed that teachers had well established their "beliefs and practices" (Allen & Lambating, 2001, p. 8) about assessment before any formal training or experiences took place (Allen & Lambating, 2001). Other studies have further made connections between a teacher's educational role models as a child and how those memories drive his or her own current assessment practices as an adult educator (Brookhart, 2004b; Shepard, 2001; Thomas & Oldfather, 1997). Specifically in music, Kotora (2001) found that vocal teachers also indicated influences that impacted their assessment practices in prior choral music education experiences.

Even when teachers have received adequate training in measurement or assessment, several studies found that classroom teachers still do not strictly adhere to the suggested practices for assessment and grading (Barnes, 1985; Brookhart, 1993; Glickman, 1993; Lomax, 1996; Manke & Loyd, 1990; Stiggins & Conklin, 1992; Stiggins, 1999). To further hinder teachers with assessment questions, textbooks in measurement in education generally may not relate to the practices of music teachers (Airasian, 1991). Although music teachers have realized the shortcomings of their current assessment practices, studies have found that few have found the means to grow

in their assessment techniques and find a reporting format that was satisfactory to all parties involved (Guskey, 2001).

Measurement. Problems with measurement issues were another challenging area identified in assessment literature (Lehman, 1998; Radocy, 1989; Rosen, 1999). This was described as a problem for any teachers who found challenge with math skills in general (Rosen, 1999). This problem was particularly noted for music teachers for two reasons (Radocy, 1989; Rosen, 1999). Rosen (1999) described the first problem of many music teachers not feeling comfortable with the mathematical skill involved in measurement even more than teachers of other subject areas, since their collegiate training would not have included many classes including these statistical concepts. The second problem for music teachers and measurement was simply that music is difficult to measure (Radocy, 1989). When statistical concepts are used with the inherent aesthetic and subjective nature of music, teachers' hesitance was credible (Lehman, 1989; Radocy, 1989; Rosen, 1999). Thus, even when music teachers do have the tools necessary to successfully complete the assessment tasks with regard to working with numbers, averages, percentages, and other calculations, problems can still exist (Lehman, 1989).

Preparation time. Assessment not only takes time during the teaching process in the classroom, assessment also takes time outside of the classroom during teachers' preparation or planning time (Boyle, 1974; Boyle & Radocy, 1987; Campbell & Evans, 2000; Gallavan, 2009a; Stiggins & Conklin, 1992). Stiggins and Conklin (1992) reported general education teachers' utilizing up to 34% of their preparatory time involved in some type of assessment-related activity. Other researchers found similar statistics in

preparation time; with some teachers spending up to 90% of the allotted preparation time available each week in assessment-oriented tasks (Campbell & Evans, 2000; Gallavan, 2009a).

Like general classroom teachers, music teachers also struggle with finding enough preparation time especially for assessment tasks (Boyle, 1974; Boyle & Radocy, 1987). As more and more responsibilities get added to an elementary music teacher's teaching load, Boyle (1974) pointed out how other responsibilities have not traditionally been additionally taken away, thus leaving very minimal assessment preparation time.

Suggestions for Improvement in Organization

The literature also cited many suggestions for improvement related to the problematic situations previously described above. Efficiency, organization, and time management were common recommendations from the literature with regard to assessment organization and are described below.

Efficiency. Finding the most efficient means of assessment was one suggestion for improvement for both general classroom and music teachers, as described in the literature. Specifically within the music education literature, Chiodo (2001) recommended the most efficient means of assessment in music was that which truly strengthens and encourages learning for the students and was well managed by the teachers. Several strategies for an efficient music assessment system were cited as follows:

1. Begin with an assessment tool that is familiar to the teacher and easy to use.

2. Document data clearly yet simply – perhaps with a seating chart or technology.

3. Imbed the assessment practices overall into the everyday learning.

4. Seek to utilize grading procedures that are appropriate to students and to the school building and district.

5. Collaborate with other music teachers.

6. Collaborate with other teachers in the building to integrate music learning – and assessment – into other classes.

7. Continually seek opportunities for simplifying assessments in the music classroom (p. 17-23).

Tuley (1985) suggested keeping evaluation or grading "simple, direct, and concrete" (p. 32). with teachers creating what was the most efficient for their teaching load or circumstances. Lavender (2000) concurred the need for efficiency and overall organization when documenting student grades, scores, or observations.

Organization. Finding a sense of orderliness was another suggestion for improvement in the literature for both general education and music education (Brophy, 2003; Lavender, 2000; MENC, 2010; Nutter, 1999). Organized record keeping, through various means, was a necessity for success in classroom assessment (MENC, 2010; Nutter, 1999). Lavender (2000) suggested specific curricular organization guides for simplifying assessment preparation tasks in *The Ultimate Music Assessment and Evaluation Kit for the General Music Teacher*. She directed music teachers to classify subsections of the curriculum and to take each section independently in a smaller piece as

to not to be so overwhelming as a broad task. For instance, a rhythm unit can seem overwhelming with regard to assessment and lacking a true starting point. However, when music teachers consider beat, tempo, specific rhythm patterns, and meter as smaller pieces within the overall rhythm umbrella term, assessment procedures or steps can be more manageable.

Another suggestion for simplification by Lavender (2000) was to start at the end, by figuring out what the end goal should be and work backwards from that goal to determine which segments or areas need to be the starting point and the sequential steps to reach the end goal. Brophy (2003) suggested an electronic means of organizing student data, which could be merged easily into reporting documents as needed.

Time management. Many authors also suggested ideas for saving time when assessing students (Bouton, 2001; Fiese & Fiese, 2001; Gallagher, 1998; Goolsby, 1999; Guskey, 1996; Hale & Green, 2009; Keenan-Takagi, 2000; May, 2001; Niebur, 1994; Nutter, 1999). Hale and Green (2009) strongly recommended that music teachers look for opportunities to assess at all times of instruction. They suggested thinking about assessment as continual, and not simply at the end of a unit or at certain parts of the school year. Bouton (2001) suggested walking around the music classroom and making observations on individual student's work during class and documenting the work immediately. Bouton suggested another timesaving suggestion by using an abbreviated record keeping system when documenting student work, such as M / E / H (mastery / emerging / needs help) or S / N / U (satisfactory / needs improvement / unsatisfactory).

In 1996, Music Educators National Conference (MENC) suggested including fewer total samples of students' work when assessing music progress (MENC, 1996b). This practice may lower reliability, but greatly reduces time requirement to complete assessments. Gallagher (1998) concurred that teachers do not assess every assignment, but balance and manage the assessments throughout instruction, relative to what was ultimately trying to be accomplished. Nutter (1999) suggested using assessment methods that evaluate the most students possible at one time if possible, in large groups rather than individually. In elementary general music specifically, seating charts have proven successful with documentation being noted right on the chart. Furthermore, music teachers can use the same performance to assess different criteria or standards (MENC, 1996b). Teachers also need to realize that not every single aspect of musical instruction needs to be assessed at every point of the school year (Fiese & Fiese, 2001).

Another suggestion with regard to minimal time in the music classroom was to "embed" (p. 42) the assessment within the typical teaching each day (Keenan-Takagi, 2000). Not only can this philosophy be a time-saver, but embedding assessments into the learning also may improve assessment quality, according to Wells (2001).

Goolsby (1999) pointed out a key factor when considering the amount of time that was required by music teachers when filling out assessment forms. He highlighted the amount of time that could actually be saved in the classroom when accurate and effective assessments are completed. Time wasted in repeating unnecessary tasks could be avoided and thus save the music teacher time. Guskey (1996) recommended classifying the overall criteria in grading into three categories: process, product, and progress.

These three decisive factors indicate firstly an interest in how the students set goals, worked, cooperated, and evolved during the learning period, or namely the process of how the students learned. Secondly, communication of the quality or score in a final project or exam would communicate a cumulative and summative achievement total. Thirdly, the final process or product could be compared to previous work achieved, and communicate effort and growth, and how the student evolved or progressed as a learner during the learning period. Guskey found that three categories included both academic and nonacademic issues described earlier in this chapter as well. May (2001) also cited these three criteria in assessment and further stated that analyzing or evaluating only one of the three criteria does not prove to be positive assessment practices.

Niebur (1994) suggested taking familiar activities and simply modifying them to include assessment qualities. For example, the students use any musical activity of singing or playing and the teacher observes and notates performance attributes in accordance with the objective. Niebur strongly recommended looking for opportunities within the established curriculum, not stopping instruction for the sole purpose of evaluation.

Practices in Summarizing Assessment Data

The third aspect of assessment as described by Stiggins (2005) was summarizing information that has been gathered and organized. Summarizing assessment data may be problematic, however, as teachers are unaware of what parts or pieces of the massive teaching repertoire summarize in preparation for reporting (Brophy, 2000). With various data gathered and multiple purposes, teachers' summarizing practices in assessment vary

widely in both general education and music education (Allen & Lambating, 2001; Bouton, 2001; Brookhart, 1994; 2004b; Cizek et al., 1995; Frary et al., 1993; Green, 2001; Olson 1989; Stiggins, 2005). I specifically wanted to inquire into what, if any, data are being summarized in elementary general music. Are teachers simply summarizing curricular or content data? Are teachers adding non-academic concepts into student summaries related to behavior? What goals do teachers have when summarizing assessment data? What are they trying to say with the summary of data? This section of the literature review, thus, targets two key ideas that identify: (a) what types of data music teachers are using when summarizing assessment, and (b) what common purpose or goals exist when summarizing assessment data in music.

Academic Factors

There are many musical factors involved within teachers' practices in summarizing assessment data (Herrold, 1991). One of the reasons why music teachers find difficulty in summarizing student work or growth was the lack of commonality with regard to exactly what should be summarized as assessment data (Bouton, 2001; Fiese & Fiese, 2001; Green, 2001; Kelsey, 2001). What academic factors are being summarized within music education assessment practices?

Indicators and standards. When determining what parts of the curriculum to summarize in preparation for reporting, it was recommended to tie assessments to curricular guidelines, indicators, or standards. Performance indicators are defined as a "specific statement that describes a performance that indicates the acquisition of required knowledge and/or skill" (p. 88). Indicators are typically designed at the local district

level to describe “specific musical behaviors” (p. 88) that indicate a student has achieved an expected level of musical proficiency or knowledge. The standards may be local or district-level standards, standards associated with an adopted music textbook series, state-level standards, or the National Standards (Brophy, 2000).

When utilizing the National Standards as an organizational guide to summarizing assessment, Achievement Standards can aid music educators with academic factors (Brophy, 2003). Each National Standard contains several Achievement Standards that were designed with measurable objectives in mind. Brophy (2003) pointed out the natural progression from each National Standard to each Achievement Standard to specific knowledge or skills that can be assessed, easily summarized, and eventually reported to others. The general academic content areas to summarize when assessing would, thus, be: singing, playing instruments, improvising and composing (both melodies and rhythms), melodic and rhythmic reading and notating, listening, and appreciation (history and cultural relationships).

Content. Within the overall umbrella of music learning, many different factors can be taught to children (Fiese & Fiese, 2001; Herrold, 1991). Singing, in general, may include pitch matching, pitch recognition and reading, rhythmic accuracy, improvisation, tone, breath support and control, interpretation, memorization, form, and historical or cultural connections (Herrold, 1991; Miller, 2005). Rhythm was another very broad category and can include steady beat, tempo, accents, visual recognition and reading, various durations, note names and symbols, rests, aural identification, syncopation, improvisation, meter groupings, time signatures, subdivisions, writing and composing,

performance, memorization, form, and historical or cultural connections (Harrison, 1983; Lavender, 2000; Miller, 2005). Melody, too, can include many different aspects of music. The staff, various clefs, sharps and flats, key signatures, intervals, aural identification, note names, ledger lines, could all fall under the concept of melody (Harrison, 1983; Lavender, 2000; Miller, 2005).

Many other similar concepts or factors also exist (Harrison, 1983; Herrold, 1991). Form may include binary, ternary, rondo, dal segno, dal capo, fine, repeat signs, repetition and contrast, and both visual and aural identification (Harrison, 1983). Harmony or texture can be considered another set of factors to consider. Intervals, triads, chords, progressions, cadences, key signatures, tonality, modality, and aural identification may all be taught within the context of harmony (Harrison, 1983; Herrold, 1991). Classroom instruments may include recorder, xylophone, and unpitched percussion, and autoharp, as well as the various other skills needed to produce a good tone and with consideration of being able to read or write the music to be played on the instruments, either alone or as a team (Brophy, 2000; Harrison, 1983; Herrold, 1991). Further areas of study may also include listening, movement, appreciation, and history studies (Brophy, 2000).

Further consideration must be given to the different means of learning music, as well (Brophy, 2000). Many of the listed curricular attributes are knowledge that needs to be cognitively acquired, defined, applied, and used as necessary. Many rhythmic, metric, melodic, harmonic, and historical elements are factual and after cognitively understanding the ideas can be used in making affective conclusions or in psychomotor

performances (Lavender, 2000; Pohl, 2000). Many musical attributes are directly associated with the performance of music through psychomotor actions, such as playing instruments, either by reading music or improvising (Brophy, 2000, Byo, 1999). Thus, one must consider both the knowledge acquisition of conceptual framework as well as the implications of this knowledge on application through musical performance.

Categories. A music teacher, like all teachers, would summarize the data that has been collected. Due to the vast number of overall musical curricular attributes as well as the different applications of musical skills and knowledge, however, categorization was suggested in the literature when summarizing assessment data (Bouton, 2001; Brophy, 2000; Lavender, 2000). Bouton (2001), Brophy (2000), and Lavender (2000) all state that when determining the summarizing procedures or steps within the assessment process, classifying all of the various music academic factors into categories may aid in the overall courses of action.

Bouton (2001) suggested sorting each musical factor into six different categories: singing, playing, reading, movement, creating, and discipline. Each of the previous academic factors can be placed into one of these categories. The category of discipline referred to any choices made by the student that did not contribute positively to the class environment or in alignment with expectations.

Lavender (2000) recommended dividing the above-mentioned factors into three major categories: rhythm, melody, and harmony. She did, however, suggest describing the particular aspects within each of these three categories with each summarization. For instance, if the melody unit of study focused on playing recorder, then that must be noted.

If the melody unit focused on notation, then that must be clearly noted. Although these three categories appear limited, these types of detailed explanations provided opportunity for each music teacher to personalize or specify details as needed, thus providing a meticulous explanation of what students had mastered (Lavender, 2000).

Further suggestions in assessment summarizing by Lavender (2000), were to divide summaries into in-class assessment data, homework data, and any other bonus data available. This delineation may later communicate to parents any significant variation with respect to student responsibility or accountability. Students were also asked to self-assess as part of the Lavender model of summarizing. Reflections on past work, as well as teacher assessment, were noted as an important aspect when seeking a thorough summarization of student achievement.

Brophy (2000) recommended the categorization of both "assessment response mode" and "learning areas" (p. 42) when summarizing assessment data. When assessments are made, Brophy stated that teachers must be cognizant of the mode with which the students are responding. The three reported modes were performing, creating, and responding. Performing was the term associated with musical tasks where students have to do something through a type of aural or visual identification, or by means of oral or kinesthetic performance. Creating was associated with skills necessitating more than just knowledge, but requiring some personal flair, preference, or choice. Responding was described as the product of a particular action or reaction of the students, such as writings, discussions, and compositions. Brophy summarized the importance of teachers

recognizing these response modes and stated that many teachers rely too heavily on performance skills only.

In addition to the response modes, learning areas were also emphasized by Brophy (2000). Although to be used only as an aid or starting point, Brophy suggested the consideration of the many academic music factors listed previously to be categorized as follows: musical skills, literacy, history and literature, analysis and preferences, and related arts and humanities. Musical skills included singing, playing, improvising, composing and arranging. Literacy encompassed the ideas of reading and writing music. History and literature was the umbrella term relating to form, styles, eras, and composers and their works. Analysis and preferences included timbre, evaluation, and the forming and communicating of predilections within various music settings. The related arts and humanities category was described as describing relationships or making connections between music and other cultures, other arts, or other disciplines (Brophy, 2000).

More important than merely forming the categories, however, was the emphasis on the relationship between categories and the appropriate means of assessment response mode. Brophy recommended music teachers using any available local, district, state, or the National Standards within the framework of "learning areas" (p. 42) with the goal of identifying appropriate assessment response modes.

Nonacademic Factors

In addition to the various academic factors, often, teachers have included nonacademic factors when figuring and summarizing grades as indicated in many studies (Allen, 2005; Allan & Lambating, 2001; Barnes, 1985; Brookhart, 1991; Brookhart,

1993; Cizek et al., 1995; Cross & Frary, 1996; Guskey, 1996; Randall & Engelhard, 2010; Wiggins, 1996). These studies have indicated that nonacademic items such as participation, attitude, effort, behavior, and punctuality, are attributes that many studies noted as being present in both general classroom and music teachers' grading practices (Brookhart, 1993, 1994; Cross & Frary, 1996; Frary et al., 1993; McMillan & Lawson, 2001; McMunn, Schenck, & McColskey, 2003; O'Connor, 1999; Ornstein, 1994; Stiggins, 2005; Stiggins et al., 1989).

More recently, two studies of both general classroom teachers and music teachers have reiterated this situation. McMunn, Schenck, and McColskey (2003) conducted a study on assessment and found that over half of the 236 elementary and secondary general classroom teachers surveyed in Florida used nonacademic factors in grading. Most recently, research in assessment in music education by Simanton (2000), Kotora (2005), and Barkley (2006), showed that traditional assessment and grading approaches in music using non-achievement criteria and efficiency as grading criterion still occur often within the profession. The findings in these recent studies reflected similar findings to the study of McCoy who found the same findings 23 years ago (Russell & Austin, 2010).

Experts in education and music education had varying opinions on whether or not the nonacademic criteria should be included in final grades (Gallagher, 1998). Several authors strongly suggested not including any nonacademic information when considering grades and assessments as they do not directly reflect any growth or advancement in actual academic learning or achievement (Bailey & McTighe, 1996; Brookhart, 1994;

Manke & Loyd, 1990; Stiggins et al., 1989; Tombari & Borich, 1999). Brookhart (1994) stated that combining other criteria into grades altered the validity of the grade. Many experts stated that teachers altered the meanings of grades when other nonacademic attributes were added in (Allen, 2005; Bailey & McTighe, 1996; Cross & Frary, 1996; Guskey, 1994; Linn & Gronlund, 2000; Nitko, 2001; Stiggins, 2001, Stumpo, 1997). Cizek (1996) explained how parents could find difficulty in "disentangling" (p. 105) a grade that was comprised of so many components, when the parents are only seeking academic growth information. Some authors thought nonacademic aspects in the classroom were difficult to measure and, thus, daunting to assess (Mierzwik, 2005). O'Connor (1999) and Ornstein (1994) noted the increase in subjectivity when including the nonacademic attributes in grading procedures. Allen (2005) pointed out the lack of meaning when parents attempt to interpret assessments that include nonacademic criteria. Brookhart (1991) went so far as to refer to the combination of such nonacademic factors in grading as "hodgepodge" (p. 36).

Other authors saw positive attributes when considering nonacademic criterion and considered such ancillary information a necessity in conveying a truly representative picture of overall student development as a learner. Hoffer and Hoffer (1987) pointed out how in music class, effort was very important and needed in many ways. Parsons (1959) referred to general education teachers' practices of including both academic and nonacademic criteria in grading as "cognitive" and "moral" (p. 304). Killian (1995) noted numerous scenarios where teachers of many different subjects justified the inclusion of such criterion. Tuley (1985) described a reporting system specifically for

music that included categories relating to such nonacademic factors and referred to them as “behavioral objectives” (p. 33). The inclusion of such factors was recommended, however, as a separate entity of the overall reporting process, not necessarily as contributing factors to overall indication of growth or progress in music learning. Gallagher (1998) posed a positive aspect of including nonacademic factors in grading when he explained how utilizing more criteria makes for a more representative or thorough picture of the students’ performances and achievements. Similarly, Airasian (1994) suggested a more wide-ranging idea of what teachers gather and include in the grading process, to get a better, more thorough, picture of where the students are with their learning and where they need to be. Walker (1998) suggested having the nonacademic criterion included on reporting tools to parents, but not figured into final grades. A separate evaluation for citizenship principles could be a positive option for music teachers to incorporate the ideas in communication with parents, yet keep the achievement and musical growth a separate entity.

Participation. One of the nonacademic areas commonly described in the literature as being included in teachers’ assessments was participation (Guskey, 2002; Kirschenbaum, Napier, & Simon, 1971; Nierman, 1997; Slavin, 2006; Walker, 1998). Stiggins (1997) described participation on students’ behalf as being a direct reflection of personalities. The outgoing and confident students participate often, and the shy, introverted students participate less often. Documentation becomes an issue with nonacademic criteria as well, as difficulties arise for all teachers when trying to define participation and document exactly who was participating and when. Especially when

considering music classes, students in music would be participating so often, that the probability of a teacher keeping students engaged and accurately documenting every detail of singing, playing, moving, and speaking was low (Mierzwik, 2005).

Attitude. Attitude was another common nonacademic area described in the literature (Airasian & Russell, 2008; Gallagher, 1998; Labuta, 1974; O'Connor, 1999; Stiggins, 2005; . Teaching and learning can be more positive experiences for teachers and students when students possess or maintain positive attitudes (Airasian & Russell, 2008; Gallagher, 1998; Labuta, 1974). The questionable issue was whether or not to use assessment and grading as a tool to encourage or even bribe good attitudes out of the students (Stiggins, 2005). Stiggins (2005) defined attitude as a feeling about someone or something, and as the root of liking or disliking that someone or something. Attitudes can vary in intensity as well, which also influences the amount of like or dislike. The literature pointed out the difficulty in defining a positive attitude and again the difficulty in measuring a positive attitude in comparison to a negative attitude (O'Connor, 1999; Stiggins, 2005). Students, just to get a better grade, again can also falsely portray this attribute.

When considered a worthy attribute to include in overall assessment, Stiggins (2005) pointed out how teachers value positive attitudes, and, thus, should include it in grading. This encompasses the idea that any tactic was acceptable to promote positive attitudes in students. Furthermore, including attitudes in grading practices can be used as "leverage" (p. 283) to maintain control and as a "reward" to the students who are successfully following rules.

Including attitudes in grading practices, however, can also be considered a negative practice. Stiggins (2005) illustrated the confusion that may arise when teachers attempt to calculate manipulations or false attitudes and simple frustration. An operational definition of a positive attitude would need to be universally accepted and accurately assessed by all teachers. Since all teachers possess different values and beliefs, this presents further challenge in equity.

Effort. Effort was another area considered nonacademic and often included in assessments (Bower & Hilgard, 1981; Brookhart, 1994; Butler & McMunn, 2006; Chase, 1999; Friedman & Frisbie, 2000; Gallagher, 1998; Lehman, 1968; Ornstein, 1994). O'Connor (1999) classified effort as "hard work" (p. 47). Although described as a highly valued attribute, O'Connor further described the difficulty to both define and measure effort. Stiggins (1997) noted the different connotations among teachers of all subjects and all levels when considering effort as well as the abilities of students to show false indicators of effort. Stiggins (2005) made clear how teachers who include effort in their assessments as being closely related if not connected to achievement and learning. Some teachers may further associate trying harder to learning more (Stiggins, 2008).

In both general education and music classrooms, the inclusion of effort in the overall grading process can be interpreted as a motivational tool to try hard and achieve more (Bower & Hilgard, 1981; Stiggins, 1997). Effort can be used as a motivator to low achieving students or to promote better behavior or actions (Allen, 1983; Oosterhof, 2001). When all other indications justify a low grade, a notable effort can improve the overall score or grade with a very positive impact (Bower & Hilgard, 1981). Our public

workforce appreciates effort and considers effort as demonstrating a necessary life skill or meeting a life goal (Stiggins, 2008).

Effort, however, was difficult if not impossible to evaluate in measurable terms (Lehman, 1968). Again, definitions vary among teachers pertaining to effort and what the precise implications are for students (May, 2001). Stiggins (2008) further pointed out that not all students are naturally assertive and that their nonassertive nature could mistakenly be considered lacking in effort. At times, the teacher may not call upon even those students who do feel comfortable with assertively raising a hand in class. Stiggins pointed out how this would cause an ambiguous situation where effort would be even more convoluted to measure or include in grading practices. Lastly, Kirschenbaum, Napier, and Simon (1971) reported that adding effort into an overall grade takes away from the intended goal of reporting on student ability. The meaning of the mark or grade loses the original intent.

Behavior. Obeying the rules and complying with teacher or school expectations was important for many reasons (Carr & Harris, 2001; Wells, 2001). Textbook authors (Butler & McMunn, 2006; Gallagher, 1998; Marzano, 2006; Phye, 1997; Stiggins, 1994) in assessment, however, have varying opinions whether or not good behavior should impact an overall grade or summation of achievement in all subjects. Grades can be used as a motivator to show up and do what was expected, as it will be in real life (Stiggins, 2005). Compliance can also lead to great student learning with a more conducive learning environment as a result of appropriate, contributing behaviors of all the students (Stiggins, 2005). Stiggins found that students can control their behavior and that can be

considered one way that students themselves can influence the outcome of their own grades. To the contrary, including behavior in the configuring of grades, may also include negative impact on assessment. If a final grade was lowered only because a student was not obeying the rules, then that final grade was not a true reflection of that student's learning or achievement (Stiggins, 2005).

Punctuality. Punctuality was another nonacademic area often included in assessments (Guskey & Bailey, 2001; Mierzwik, 2005; O'Connor, 1999; Stiggins, 1994). Two aspects of punctuality were discussed in the literature with regard to both general and music classrooms: (a) students being on time to class as responsible citizens, and (b) work handed in on time as assigned (Guskey & Bailey, 2001). When punctuality was intended to have an impact on overall grading, documentation can be very problematic. Mierzwik (2005) noted the organizational difficulty that accompanies keeping track of assignments handed in late, at random times past the deadline, as well as entering those late scores.

O'Connor (1999) noted the problems associated with penalties for submitting late work. When points are deducted with tardiness, students lack the initiative to complete the work at all, realizing that the work will not receive many points at all with the deductions. Thus, if teachers truly want quality work to be completed, thought must accompany the message that is sent to students with late work. From an anonymous source, O'Connor quoted, "It is best to do it right and on time, but it is better to do it right and late than the reverse" (p. 50). O'Connor found that teachers ultimately hope that

students turn in work on time, but dealing with the tardiness in ways other than grading penalties was recommended.

Teachers in both general classrooms and music classrooms are summarizing data in multiple ways and with multiple ideas (Allen & Lambating, 2001; Brookhart, 1994; 2004b; Cizek et al., 1995; Frary et al., 1993; Olson 1989; Stiggins, 2005). Overall, one clear means of organizing assessment data was not gleaned from the literature in either general education or music education. The overall theme that was clear from the literature was the discrepancies among all teachers as to how to effectively summarize the data most effectively. Nor was there a clear procedure or rationale for exactly what factors – academic or nonacademic – should be summarized in music assessments. From the review of literature, thus, a clear gap in practices in summarizing assessment data was evident.

Practices in Reporting Assessment Data

Reporting student progress was an important aspect of overall assessment practices (Stiggins, 2005). Reporting practices, in all areas of education, are to inform students and parents of progress, growth, and learning by sharing assessment data (Nutter, 1999). Teachers have utilized many different reporting methods to communicate student learning. The ultimate goal was to inform parents on the progress of their child's learning and growth (MENC, 1996b). The most prevalent means of reporting assessment data as reported in the literature were grades, report cards, and conferences, either used individually or in combination (Bailey & McTighe, 1996; MENC, 1996b). This section of the literature review will investigate each of these practices.

Grades

One format used by general education and music classroom teachers was grades (Harrison, 1983). Many teachers use the terms assessment and grading interchangeably (Asmus, 1999). Grading was defined as assigning a score or a letter grade to student work (Fautley, 2010). Brophy (2000), however, noted many options for schools to document grades on student progress: letter grades of A through F; ratings of excellent, satisfactory, needs improvement, unsatisfactory; and proficiency levels of advanced, proficient, basic, and needs improvement.

Studies in grading practices of schools relating to assessing children have indicated much dissatisfaction with grading patterns and reform efforts (Labuta, 1974). These studies also indicated the frequent use of teachers using personal preferences to grade and including many various achievement and nonachievement factors in assigning grades (Cross & Frary, 1999; Guskey, 1996; Olson, 1995; Tyack & Tobin, 1994).

Report cards

Another tool used in elementary and secondary schools for reporting student learning was a report card (Allen, 2005; Bailey & McTighe, 1996; Harrison, 1983; Lake & Kafka, 1996; Shuster, Lynch & Polson-Lorczak, 1996). The report card was described in the literature as being the most common form of reporting or sharing assessment information with parents (Allen, 2005; Bailey & McTighe, 1996; Harrison, 1983; Lake & Kafka, 1996; Shuster et al., 1996). Harrison (1983) described a report card as a tool to invite parents to become involved in the music program as they serve as “commentary” (p. 313) from the teacher to the parents regarding their student and the program. Nutter

(1999) reported that schools send home report cards in music as often as every six weeks or as rarely as twice per year.

Tuley (1985), in an attempt to share a model of evaluation reporting format to be "understood and used by many" (p. 32), suggested five categories to include in an elementary music report card: (a) music skills, (b) concepts, (c) participation, (d) conduct, and (e) unique contribution or problem. Secondly, he suggested keeping the report card simple and to the point for the most convenience to the music teacher. Each of the five categories had very easy rating scales with several options for the music teacher to check. Tuley further recommended aligning with individual district or building objectives to guide alterations of specific content within each of the five categories.

Nutter (1999) described a report card format specific to elementary general music where teachers do not summarize all music skills, achievements and behaviors in one overall grade, but rather into several categories and highlighting strengths and weaknesses within each area. Chase (1999), however, pointed out how one format of report card cannot meet the needs of all teaching situations. Guidelines are just that – guidelines. When designing report cards, Chase suggested getting input from many different parties and seeking information that these parties want to glean from the tool. Teachers should also consider the achievement aspects that should be communicated to parents and determine how best to use the report card to do so, strongly considering the format that would be best comprehended by parents.

There are many advantages and disadvantages of using report cards (Chase, 1999; Harrison, 1983). The advantages were the overall means to convey student progress, clarity to meet parents' needs, and comments to reinforce grading data (Chase, 1999). The disadvantages were listed as minimizing student work to a single mark, little information regarding proficiency levels, and overall overgeneralizations that lack deep meaning (Carr & Harris, 2001; Chase, 1999). Carr and Harris (2001) described how report cards may be either extreme of helpfulness, depending on the content. Too little information can be interpreted as sparse and uncaring. Too much information on a report card can be overkill with unnecessary details that parents do not need.

The following attributes were listed as positive traits of report cards: easy to understand descriptors, connection to standards, produced electronically--not by hand, inclusion of citizenship qualities separately, adjustable with different times of the year, containing clear language without jargon, addition of written comments, inclusion of sub-skills of each content area, inclusion of definitions of any unfamiliar terms, and reflective of the actual assessment tools used in the classroom (Carr & Harris, 2001). Power and Chandler (1998) suggested adding attachments to a report card format, especially if a parent-teacher conference was not included in the scenario. Attaching a cover letter or other attachments of documentation or explanation could prove beneficial.

Guskey (2001) noted the importance of clarifying the purpose of each reporting tool and pointed out the common error of many educators developing report cards without consideration of the purpose. Guskey recommended considering what information was to be communicated, to whom it will be communicated, and how the

information will further be used. Stiggins (2005) emphasized two aspects of report cards that must be present: timeliness, and clarity. Waiting too long to inform parents of students learning situations can automatically blur the main points of concern or strength. When parents are confused by information shared, they will not understand the true message about their child's growth and learning development (Stiggins, 2005; MENC, 1996b). Regardless of how the report card is formatted, the tool must communicate how well the students are doing in the music classroom (MENC, 1996b).

Conferences.

Conferences between teachers and parents are another common method of reporting student progress in both general and music education (Bailey & McTighe, 1996). Harrison (1983) reported that teachers should make "frequent" (Harrison, 1983, p. 313) communication efforts with parents and found that teachers conducted conferences to let families know what was happening in music and how their child was functioning. Stiggins (2005) pointed out the many benefits to conferences as a means to communicate assessment data. Verbally, teachers can give a much clearer idea of what was or is actually going on in the classroom with opportunities for questions, new topics, and personal sharing.

Parent-teacher conferences are an opportunity for communication to be "interactive" (p. 103) and give parents the opportunity for parents to ask questions regarding teacher comments or grading (Guskey, 2002). The format allows for clear communication through truthful words where feedback can be immediate and can alleviate any confusion or miscommunication (Bailey & McTighe, 1996). Each

conference was also considered “individualized” (p. 103) as teachers are talking primarily about only one child’s strengths, weaknesses, and accomplishments (Guskey, 2002). Howe and Simmons (2005) suggested starting conferences with positive and friendly comments, accentuating the positive attributes of the student and following up each comment with documented examples of student work. Parents can communicate with teachers about ways to support learning at home as they interpret the areas that need more support when learning (Stiggins, 2005). Teachers should also listen sincerely to parents’ concerns and questions (Howe & Simmons, 2005).

Teachers’ Motivations for Assessment

Another important aspect of this research study was the motivations, purpose, or reasons relating to the decisions that teachers are making with respect to assessment (Airasian, 2000; Anderson, 2003; Colwell, 1974; Foley, 2001; Hamann, 2001). Bailey and McTighe (1996) stated that teachers “have used grades and reports for many purposes” (p. 20). Many different authors within the research literature have noted many different purposes for assessment and for grading (Airasian, 2000; Airasian & Russell, 2008; Allen & Lambating, 2001; Anderson, 2003; Bailey & McTighe, 1996; Colwell, 1974; Foley, 2001; Gallagher, 1998; Gredler, 1999; Hamann, 2001, Linn & Gronlund, 2000; Nitko, 2001; O’Connor, 1999; Oosterhof, 2001; Stiggins, 2001; Taylor, 2003). The different rationales of assessments, grades and reports were described throughout the literature as (a) providing incentives to motivate students; (b) evaluating progress; (c) identifying and labeling students for awards, honors, and special programs; (d) evaluating teachers and programs, and (e) communicating progress to others (Airasian, 1994; Austin

& McCann, 1992; Farrell, 1997; Hoffer, 1993; Marzano, 2000; Silberman, 1970; Strickland & Strickland, 1998; Taylor, 2003; Terwilliger, 1971; Wrinkle, 1947).

Many other authors agreed on one main purpose of assessment in music as the gathering of information about what students are learning in music while clearly distinguishing a relationship between learning objectives, teaching, and assessment (Asmus, 1999; Chiodo, Frakes, MacLeod, Pagel, Shuler, Thompson, & Watson, 1998; Duncan, 2009; Lehman, 1992; McTighe & Ferrara, 1998; Radocy, 1995; Swanwick, 1998). Another factor revealed in the research literature as a role of assessment specifically in music education was the positive learning incentives provided to students as they reflect on what they have learned and have the opportunity to value their progress (Farrell, 1997; Hill, 1999; McClung, 1996; Niebur, 1997). Both Asmus (1999) and Cope (1999) shared more insight on the idea that assessments played a larger role than simply contributing numbers to figure a final grade. Rather, assessments assisted in determining whether progress was being made toward the objectives. Shuler (1996a) extended the idea that assessment also made it possible for teachers to determine whether their teaching was being productive in helping students learn.

In addition to the role or purpose of assessment in music education, many different motivators are present when music teachers are contemplating assessment (Talley, 2005). Although most studies indicated that music teachers did not agree on any one particular factor as a single motivator for assessment in the music classroom, many studies indicated that the majority of teachers do value assessments as an important part of their job (Barkley, 2006; Brummett, 1993; McClung, 1996; Monroe, 1995; Niebur,

1997; Nightingale-Abell, 1993; Talley, 2005). Many studies indicated that monitoring student growth, progress, and achievement was the primary motivator for assessment in elementary music (Brophy, 2000; Niebur, 1997; Nightingale-Abell, 1993; Talley, 2005). Other studies found that teachers were motivated to assess students in music class simply because of the relationship between assessment and instructional goals within the overall music curriculum (Brophy, 2000; Fiese & Fiese, 2001; Niebur, 1997; Nightingale-Abell, 1993; Talley, 2005).

Finally, accountability has been found to be a strong motivator to assess in music classrooms, as teachers are often defending the importance of the music program in times of budgetary cuts (Brophy, 2000; Farrell, 1997; Niebur, 1997; Schultz, 2002; Talley, 2005). Talley (2005) pointed out the importance of substantiating students' learning evidence when administrators are considering cutting back in programming at schools. Assessments do indicate to all interested parties that learning has occurred (Duncan, 2009). Similarly, Colwell (1995) noted how political circumstances have played a role in music teachers' increased interest in assessment. As budgets continue to be an issue in most states and school districts, Colwell strongly advised all music educators to improve upon all facets of music education, most specifically, assessment practices. If the budgetary decisions are truly data-driven decisions, then student growth and new assessment measures in music must be a part of this scenario in a positive way.

Table 1. *Research Studies in Music Teachers' Practices with Assessment.*

Level	Year	Author	Findings
Elementary	1979	Hartwell	Inadequate release or preparation time to accommodate assessment (and other) tasks.
Elementary	1986	Carter	Assessments based on effort and participation. Recommendations for increased objectivity and specific grading criterion.
Elementary	1988	Rasor	Teachers utilize paper/pencil and verbal testing practices. Most teachers use letter grade or satisfactory/unsatisfactory. Most often assessed criteria include singing, instruments, movement, and attendance.
Elementary	1990	Miller	Affective responsiveness to music can be observed and charted.
Elementary	1993	Brummett	Teachers found framework of process-oriented portfolios as valuable way to document student achievement and growth.
Elementary	1993	Nightingale-Abell	Teachers lacked training in assessment practices. Teachers working with large number of students with minimal contact time. Teachers favored informal assessments and performance assessments.
Elementary	1997	Anderson-Nickel	Differences existed in assessment practices between novice and experienced teachers. More experienced teachers reported more organization and consistency in assessment practices.
Elementary	1997	Niebur	Teachers sought improvements with assessment. Teachers' first priority musical experiences. Teachers' motivation for assessment varied.

(Table Continues)

Level	Year	Author	Findings
Elementary	1997	Shih	Teachers use assessment for instructional planning and student placement. Teachers focus more on teaching than assessing. Curricular alignment not highly relative to assessment practices. Many curricular variations in teaching and subjectivity found acceptable.
Elementary	2004	Hepworth-Osiowy	Teachers believe assessment is valuable. Teachers use variety of tools and strategies to assess. Difficulties common and lead to feelings of inadequacy.
Elementary	2005	Talley	Teachers viewed assessment as means of legitimizing and providing validity to music in the curriculum. Very little use of commercial achievement or aptitude tests. Most assessments informal.
Elementary	2006	Barkley	Teachers favored informal assessments and performance assessments. Inhibitors to assessments were time, training, schedule, and resources. Assessments based on effort and participation.
Elementary	2011	Salvador	Case studies and cross-case analysis found a variety of assessment methods throughout school year with primary purpose of informing instruction. Hindrances to assessment also identified as large number of students, lack of time, and lack of support.
Secondary	2010	Russell & Austin	Teachers lacked guidance. Practices varied greatly. Grading based on achievement and nonachievement factors. Schedule and number of students did not factor.

(Table Continues)

Level	Year	Author	Findings
Secondary	1995	Monroe	Teachers' ideal assessments and actual assessments differed greatly. Teachers and principals believed assessment should be performance based and measurable.
Secondary	1988	McCoy	Teachers had much variation in grading criteria. Teachers had lack of understanding in determining grades. Nonacademic criteria included in majority of music grades.
Choral	1986	Weymuth	Author-developed Choral Music Achievement Test considered useful tool for evaluation of important nonperformance objectives in Missouri.
Choral	1996	McClung	Teachers and principals believed assessment should be tied to objectives. Teachers supported considering participation and attitude in grading. Assessment in music impacted public perception of music. Teachers were content with current practices.
Choral	2002	Tracy	Teacher assessment practices based on personal priorities.
Choral	2005	Kotora	Teachers lacked guidance in assessment practices. Teachers found challenge in managing students while assessing. Nonacademic factors utilized in grading. Parent and student apathy toward music found.
Instrumental	1999	Hill	Teachers believed assessment to be important. Teachers believed grades to be motivating and a discipline tool.
Instrumental	2000	Simanton	Assessment formats vary. Assessments based on nonacademic factors. Teachers content with current practices. Smaller bands closer to best practice. Teachers with graduate degrees closer to best practice.

(Table Continues)

Level	Year	Author	Findings
Instrumental	2001	Hanzlik	Situational factors impede assessment efforts
Instrumental	2001	McCreary	Teachers use traditional assessment tools, with most not utilizing alternative methods. Teachers included non-musical criteria in calculating grades.
Instrumental	2002	Sears	Teachers formally assess students at different increments throughout year and with variety of assessment tools.
Instrumental	2006	Kancianic	Teacher assessment practices based on personal priorities.
Instrumental	2006	Sherman	Some consistency in assessment strategies, yet many irregularities still exist. Grades calculated in various ways.
Instrumental	2009	Duncan	String teachers commonly assess informally. Success or string programs correlate with particular assessment practices.

Teachers' Practices with Assessment in Music Education

A review of the research specifically in music education revealed that the majority of the research conducted on assessment in music education included surveys of assessment and grading practices. The assessment research in elementary general music, secondary choral, and secondary instrumental was summarized in Table 1. The studies are presented in Table 1 in chronological order, categorized by elementary and secondary areas.

Elementary Music Research Studies

The studies from Table 1 illustrated the lack of a common format or guidance in assessing students in all music settings. As noted, many research studies focused specifically on assessment in elementary general music. Hartwell (1979) conducted survey research and found an affect between evaluation of students and class size. Although this study focused on many different areas of elementary music education, assessment was a noted area of concern. Of the 436 respondents in this study of Ohio music teachers, many responded that due to the increased number of students, there simply was not enough time to work on evaluative practices within the schedule.

Carter (1986) administered a survey to elementary general music teachers in Oklahoma. Of the 461 respondents, the researcher summarized strengths and weaknesses of elementary music programs in general. In addition to other findings, assessment practices were highlighted as in need of improvement. Specifically, the elementary music teachers surveyed indicated deficient objective testing procedures. Recommendations from the study included more consistent and objective testing practices and less reliance on classroom behaviors and participation when factoring final grades.

In a 1987 study, Rasor (1988) investigated the general practices of elementary music programs in Ohio. Many different systematic criteria were studied including evaluation practices. Of the 655 participants in this study, fewer than 10% indicated using standardized tests. The majority of respondents indicated that they used paper and pencil tests and questioning techniques in the classroom. The content of the evaluations

were primarily singing, performance (instrumental), movement, and attendance. Reporting practices for a majority of participants included letter grades or 2-criteria means--either pass / fail or satisfactory / unsatisfactory.

Brummett (1993) found portfolios to be viable and valuable sources of assessment for elementary music teachers. In this study, Brummett sought to provide a descriptive and interpretive narrative of two teachers' context for process-oriented student evaluation. The two Missouri music teachers who took part in this qualitative study utilized portfolios that contained students' checklists, tapes, quizzes, journal notes, reflections and projects over time. These frameworks, developed by the research and adapted by each teacher, provided quality alternative modes of assessment for these two elementary general music teachers. The study also explored how the teachers utilized and modified the framework in sixth grade general music and also their reactions after implementing them.

Observations, interviews, and document review by Brummett (1993) over eight months took place. Conclusions included student preferences of certain musical performances and creative efforts where self-discipline, self-direction, and self-evaluation were the focus of the classroom environment. The two teachers found many positive attributes with the framework including curricular flexibility, adaptability to classes and programs, and emphases on the instructional cycle and continual student independence. The portfolios that were the product of the study were found to be a positive attribute by both the teachers and she students.

In another study of elementary general music, Nightingale-Abell (1993) conducted a qualitative study of three veteran music teachers. Following lengthy observations, video reflections, document review, and interviews, conclusions were made that the teachers utilized much informal and formative evaluation strategies. The teachers in this study utilized many various informal assessment strategies, including many involving observation and even what the research referred to as "mental record keeping" (Nightingale-Abell, 1993, p. 193). This study also revealed that the quality of evaluations was based on the amount of planning time allotted to each teacher.

Anderson-Nickel conducted a survey in 1997 regarding elementary music teachers' experience in teaching music using qualitative research methods to investigate the differences between novice, advanced beginners, competent, proficient, and expert music educators. Her research study included interviews and observations of 12 music teachers of varying experience levels. Although this research was not primarily investigating assessment, one of the criteria utilized by Anderson-Nickel was evaluation and assessment. With respect to these criteria, this researcher found differences in evaluation between novices and experts. The teachers with more experience kept records of data collected through observations by using seating charts or checklists for student performances and student participation where the less experienced teachers relied on memory for record keeping. The more experienced teachers were also found to have more consistency in assigning grades. Further findings showed that all participants were somewhat confident with diagnostic activities and that the more experienced teachers

were able to work with assessment activities in the classroom without losing management control.

Niebur (1997) investigated four elementary music teachers regarding standards and assessment. The two-year-long qualitative study took place at the same time the National Standards were being established. These four teachers, who had all completed a master's course in assessment and standards, sought improvement options with regard to assessment, considering numerous assessment options. Conclusions were made that assessment was not, however, the first priority of these teachers, but rather providing quality musical experiences was the main concern.

Curriculum alignment of fifth grade general music in central Texas was the focus of Shih's (1997) study. This study specifically investigated the relationship between the objectives within the state curriculum and the objectives that the music teachers actually taught and assessed. Shih utilized a survey instrument and also interviewed 15 percent of the participants. Shih found that state standards were taught and assessed but not necessarily on an equal basis. The survey also asked teachers about which state objectives were assessed on a regular basis. A majority of the respondents indicated they assessed singing, listening, movement, and notation objectives. Teachers also responded that about one fifth of the state mandated objectives were not assessed. The assessment method most often utilized was observation during group performances, and observation of individual performance was also used.

Further responses indicated variation among schools included different textbooks, different teacher qualifications, different preparation time, and different administrative

policies. Overall, Shih found a low degree of curricular alignment. The research summary indicated three factors associated with the low alignment. Firstly, assessment and instruction are not clearly reliant on any district or state guidelines for music, like they are in the general classroom. Secondly, limited student contact time was also noted as another inhibitor. Lastly, the respondents' answers denoted a general disinterest for assessment. Considering these three findings, Shih concluded that the overall lack of an "assessment system" to evaluate learning and achievement was the main reason why music teachers do not assess more often and more systematically in elementary general music.

Hepworth-Osiowy (2004) surveyed 190 elementary music teachers in Canada regarding practices and perceptions in assessment. Hepworth-Osiowy reported that the teachers who responded to her survey valued assessment and utilized a multitude of assessment strategies and practices to most accurately assess their own students, either on-going in each class or less often and irregularly. When asked to reflect on assessment factors, a majority of participants identified negative aspects and did not identify any "positive factors impacting… assessment practices" (p. 107).

As referenced earlier, Talley (2005) surveyed 35 elementary general music teachers in Michigan seeking the frequency, methods, objectives, and applications of assessment. Data collection determined very infrequent informal assessments to be most common featuring observation, original tests and worksheets, and group performances. Very little commercial use of achievement or aptitude tests was found. Content areas that were most commonly evaluated were beat, singing, pitch, rhythm, recorders, note names,

instrument families, and music literacy. This researcher also concluded that the music teachers reported two sources of motivation to assess. One motivator was reported as justification for music programs and a second motivation aimed to "adapt and individualize instruction" (Talley, 2005, p. 60).

Barkley (2006) investigated the current practices of elementary general music teachers with regard to assessment strategies, frequencies, influences and attitudes, as well as the National Standards for Music Education. This descriptive research study included 255 elementary music teachers in Michigan. Data collection indicated teachers' use of the National Standards to some degree as criterion for assessments, with personal preferences as the determining factor as to which standards were most frequently assessed. Observation was the most common assessment technique utilized, without, however, noted details of how the observations were documented. Portfolios were found to be the least utilized assessment tool among the survey respondents. Furthermore, inhibitors to assessment were found to be time, resources, schedule and training. The teachers in this study also indicated familiarity with the National Standards for Music Education and felt that these standards should be a part of assessment practices. The majority, however, indicated that effort and participation were the "most important" (Barkley, 2006, p. 52) criterion when reporting student progress.

Secondary Music Research Studies

Although not a part of this research project, many secondary studies in music assessment were also reviewed. The research studies that have occurred in secondary

music are also included in Table 1. The secondary research studies revealed a broad range and variety of attitudes and practices.

Summary of Research Studies in Music Assessment

Overall, practices in music education assessment and grading have not changed significantly in many years (Russell & Austin, 2010), nor have any new methods or approaches been devised for common adoption or implementation in elementary (or secondary) music education classrooms, as reported in the studies found in Table 1.

Summary

This review of literature was aligned with Stiggins' (2005) assessment categories, and the research questions of this study. Additionally, this literature review briefly examined the history of assessment in education, and further investigated teachers' motivations for assessment and described the related research studies with respect to music teachers' practices with assessment.

The history of educational assessment in both general education and music education can be summarized as teachers who have continually striven for continued improvement with an abundance of barriers present. Assessment has become a target area within education as a means for improving teaching and learning in times of educational crises.

A wide variety of assessment practices was found as both general classroom and music teachers assess with numerous and varying techniques and tools. Although some gathering practices are more common than others, neither educators nor experts in any subjects agree upon any common assessment practices. A clear gap of common or

consistent processes, methods, or practices was evident. This gap of commonality justified the need for this study as well as a basis for further research in this area.

Organizational practices revealed prevalent problems associated and reported with assessment in both general and music classrooms with high numbers of students, lack of instruction time, limited training, measurement challenges, and lack of preparation or planning time. Several suggestions for improvement as described in the literature were ways to improve organization and communication through efficiency and effectiveness. This information supported the research goal of providing further suggestions for improvement for organizing elementary general music assessment data in Iowa.

Teachers' efforts to summarize assessment data were also reported as problematic with dilemmas of exactly what data to include when summarizing. Both academic and nonacademic factors were considered in teachers' summaries. Recommendations to keep nonacademic factors in a separate citizenship report were made. Sharing and reporting of assessment information to parents and other interested parties were found in the form of report cards, through conferences, and using grades. The fact that the review of this literature revealed numerous different influences on summarizing assessment data further supported the need to investigate what Iowa elementary general music teachers are doing.

The motivation for assessment in both general education and music classroom was found to vary greatly. Teachers indicated throughout the literature that assessment was used as incentives for students, to evaluate student progress, to evaluate teachers and programs, to identify both excellent and struggling students, to connect to instructional goals, and for accountability purposes.

The many music research studies included in this literature review also reiterated the lack of commonality in assessment practices of gathering, organizing, summarizing and sharing. These facts reinforced the need for similar questioning of the respondents in this study. This literature review also served as a guide in providing samples and models for surveys and also specific questions for this project's survey instrument.

In addition to identifying the gaps in practices and the lack of research within the state of Iowa, this literature review has provided further insight to the overall literature with respect to elementary general music assessment practices. This literature review also led me to this research study by situating this study within the broader scholarly community.

CHAPTER 3

METHODOLOGY

The research design for this inquiry was a mixed methodology of surveys and focus groups. The quantitative and qualitative procedures for collecting and analyzing data are discussed in the chapter. The foundation of the methodology was the three research questions.

1. What are the features of the classroom assessment practices currently implemented to:
    a. gather evidence (assessment data) of student learning in elementary general music in Iowa?
    b. organize student learning data in elementary general music in Iowa?
    c. summarize assessment data in elementary general music in Iowa?
    d. report assessment data in elementary general music in Iowa?
2. What demographics, or teacher characteristics, influence or relate to classrooms assessment practices of gathering, organizing, summarizing, and reporting assessment data in elementary general music in Iowa?
3. What are teacher beliefs or opinions related to assessment and reporting in elementary general music in Iowa?

The chapter is comprised of sections that include or describe the rationale and design, population, survey sample selection, survey instrument development, pilot study, survey instrument content and sources, procedure, survey data analysis, qualitative focus groups, focus group sample selection, and qualitative data analysis.

Rationale and Design

This section will discuss the rationale and design of this mixed methodology study. Both the rationale and the design relate to Stiggins' (2005) four general assessment categories of gathering, organizing, summarizing, and reporting, and the research questions stated above.

Rationale

Assessment is an essential piece of instruction for all teachers (Airasian, 2000; Brophy, 2000; Haladyna, 1999; Lehman, 1968; Radocy, 1989). Education and music education have both received criticism with regard to assessment and, as indicated in the literature review, many of the problematic situations in assessment have existed for many years (Boyle, 1974; MENC, 1996b; Shuler, 1996b). Although teachers were reported as having an awareness of music assessment and related problems, specific procedures or strategies for improvement were not clarified (Davidson, 1995).

Since the inception of the National Standards in Music (Consortium of National Arts Education Associations, 1994), school districts and individual school buildings have set goals to determine the extent to which students are achieving the standards (Wells, 2001). Even as the standards were described as "setting the stage" (Lehman, 2000, p. 8) for assessment in 1994, many researchers emphasized the problems and the many variations associated with assessment in music (Barkley, 2006; Goolsby, 1999; Kancianic, 2006; Lehman, 2000; McCoy, 1988; Russell & Austin, 2010; Shih, 1997; Simanton, 2000; Talley, 2005).

Most recently expert authors and researchers have emphasized the importance of assessment, while, at the same time, have identified the lack of any common methods or consensus on acceptable approaches to assessment in music. The Music Educators National Conference (MENC, now known as The National Association for Music Education, NAfME) published the *Spotlight on Assessment in Music Education* (2001), which contained 31 articles, most of which described the varying ways music teachers of all levels assess different musical skills. In 2002, *The New Handbook of Research on Music Teaching and Learning* identified assessment as one of the "more important issues in education" (Colwell, 2002, p. 194), yet no prescribed methods were noted. In both 2007 and 2009, the University of Florida hosted music education symposia on assessment. Experts in music education assessment such as Richard Colwell, Paul Lehman, and Tim Brophy led meetings and published proceedings relative to assessment (Brophy, 2008; Brophy, 2010). At the 2009 Symposium, Colwell (2010) expressed the difficulty of assessing music when he summarized "though many outcomes may be hard to capture on a test, that does not mean that the teacher ignores... them" (p. 16). Again, these experts in music education have not found data supporting consistency in elementary general music assessment and thus the main rationale for this research is this problematic gap. Seeking what currently exists in assessment in music education in Iowa, as was the purpose of this research, can further provide a foundation on which to build recommendations and steps for more quality and excellence in music assessments.

Design

The research design for this descriptive study was mixed methodology using both quantitative and qualitative data. Surveys and group interviews provided a means to collecting both quantitative and qualitative data within Stiggins' (2005) four assessment categories of gathering, organizing, summarizing, and reporting the information, relative to the first research question. The data was further analyzed to seek relationships between similar demographic responders (second research question). Furthermore, beliefs and opinions were sought through further survey and focus group questions. The analysis and findings of this research study has provided rich descriptions of the practices and perceptions relating to assessment in elementary general music in Iowa.

Population

The target population for this study is the elementary general music teachers in the 872 Iowa public and private schools that house elementary grades during the 2012-2013 school year. This population, or the sampling frame, of elementary schools in Iowa was provided by the Iowa Department of Education, Bureau of Planning, Research, and Evaluation. The list consists of 729 public schools and 143 private elementary schools from all 99 counties in Iowa, divided into nine Area Education Agency centers.

The sampling frame was restricted to educational buildings that house students in any elementary grades of Pre-Kindergarten through sixth grade. The sampling frame did not include buildings specifically labeled intermediate schools or middle schools. Although some intermediate or middle schools contain students in the target grades, these buildings tend to teach music in ensembles and not general music classroom settings. All

buildings with students in the elementary (PK through sixth) grades were a part of the sampling frame, even if the school building housed additional grades, such as a Pre-Kindergarten through eighth grade school or a Pre-Kindergarten through twelfth grade building.

Quantitative Research: Survey

The first portion of the research included the collection of quantitative data through surveys of elementary general music teachers in Iowa. The quantitative aspects of this study included communication with identifying a sample of the population, the creation of a survey instrument, the distribution of the survey instrument, and analyzing the collected data.

Procedure for Selection of the Quantitative Sample

To select the sample for this research study survey, a master list of all public and private elementary schools in Iowa was obtained from the Iowa Department of Education in Des Moines. The target population for the survey research was 872 public and private elementary school buildings. From this population of elementary general music teachers in Iowa, a representative sample was created. To ensure that the sample selected for this study had a proportional number of schools from throughout all regions of Iowa, random stratified sampling was done, with each of the nine AEA regions utilized as the relevant stratification categories. This reduced the potential for bias in the selection of participants to be included in this sample while maintaining representation of the entire population of music teachers in Iowa. This process also reduced possibilities of threats to external validity and the chance of sampling error.

To create a stratified random sample, the entire population of 872 schools were organized by AEA region, with all of the schools within each stratum listed alphabetically by school name. Each school was then assigned a consecutive number. The sample size for this research study was set at n=315, which provided adequate data for a representative sample as well as a manageable number for adequate study organization. To achieve a sample of 315 music teachers, 35 schools were chosen from each AEA strata using a random number generator. Communication took place with every elementary general music teacher at each of the 315 schools in the sample.

In addition to the 315 randomly selected schools, all of the schools in the county and two surrounding counties of the researcher were purposefully selected if not already chosen by random. Many of these respondents were also asked to participate in the qualitative portion of this research project and also needed to be a part of the quantitative data collection. All of the focus group participants willingly indicated participation in the survey.

Many steps were put in place to ensure both a representative sample and external validity. Pre-announcements were sent to the sample, communicating the upcoming study and survey details. Respondents were able to locate the respondent-friendly survey instrument online. Reminder e-mails were sent to music teachers from the sample that had not responded within two weeks, all in an effort to achieve a high response rate. Appreciation communications were also sent to those who have completed the survey.

Development of Instrumentation

The survey instrument created for this study grew out of information gathered from a pilot study, the literature review, surveys from other research articles, and advice from a survey expert. These various influences aided in the development of the final version of the survey instrument, entitled *Elementary General Music Survey on Assessment in Iowa.* The specific sources for each of the questions are found in Table 2, with brief descriptions as well.

Pilot study. A pilot survey was sent via e-mail to 25 elementary general music teachers who had recently been participants at a music education workshop and indicated informal willingness to respond to a pilot study. The pilot survey was comprised of 10 general questions with regard to assessment in elementary general music, based on years of elementary general music instruction by this author, personal inquiries in survey development, and a beginning literature review. The pilot survey is attached as Appendix A. Responses were received from 22 music teachers, an 88% response rate. The respondents ranged from 0 (first year teacher) to 40 years of music teaching experience.

In addition to general demographic questions, the questions on the pilot survey instrument asked questions aimed at discovering music teachers' practices relating to assessment. Teachers responses regarding summarizing and reporting assessment data were clear from the piloted questions and multiple choice answers provided. The areas of gathering and organizing, however, were not clearly described with the piloted format. It became clear that more questions were needed, specifically targeting the gathering and organizing components of the research questions.

In addition to adding more questions, other improvements were made to the survey instrument with consideration of closed question responses and open-ended answers. The pilot survey question with the richest, most in-depth responses was the open-ended question regarding challenges in assessment (question number 9). This mode of questioning was replicated several times in the final survey instrument to attain stronger responses. Although this pilot question did acquire responses that were very personalized, detailed, and specific, it was also noted that only 15 of the 22 respondents answered this question, which could indicate, in general, a respondent's preference for faster, easier answers and a desire to complete the survey quickly and with ease, avoiding questions that would require a more invested and lengthy contribution. Thus, the final survey instrument is not overloaded with these open-ended questions, but a balance of several types of response options.

In addition to the open-ended question, the pilot survey contained several questions with multiple choices for the respondent to choose. Unfortunately, the data collected from these questions did not indicate respondents' need to express other possible responses that were not present within the given list, but rather limited the respondents to only the choices and attributes that were present. These responses given in the pilot, limited the respondents' options, and a majority of the respondents chose all of the options, as they were instructed to choose "all that apply" (question number 10). The change from this type of question in the final survey was to provide an opportunity for the respondent to add any comments, additional information, or specific rationale

relating to the question, beyond just the yes or no indication for each choice or attribute listed.

Literature review. The literature review, summarized above in Chapter 2, also provided resources for the compilation of survey questions. Each section of the literature review focused on one of the four main components of assessment, as outlined by Stiggins (2005): gathering, organizing, summarizing, and sharing and reporting assessment data. As each of these sections of the literature review was researched and written, specific characteristics and common themes arose as the sections grew in breadth and depth. These commonalities within each main section of the literature review, then, became natural guidelines for either additional questions, or options for responses to the survey questions.

With reference to gathering practices, many sources (Airasian, 1994; Berman, 2008; Chittenden, 1991; Farrell, 1997; Foley, 2001; Gronlund & Linn, 1990; Hill, 2008; Johnson et al., 2002; Mierzwik, 2005; Salmon-Cox, 1981; Slavin, 2003; Stiggins, 1994; Stiggins, 2005; Taylor, 2003; Trice, 2000) noted the most often utilized means of gathering assessment data at the elementary level as examinations, projects, assignments and homework, observations, performances, portfolios, and rubrics. These items then became the natural multiple-choice responses for questions relating to gathering practices.

With regard to the practices in organizing assessment data, many sources (Airasian, 2000; Brophy, 2003; Carter, 1984; Colwell, 1974; Dorr-Bremme & Herman, 1986; Hamann, 2001; Lehman, 1968; Marzano, 2000; Russell & Austin, 2010; Walker,

1998) identified the most prominent problems in music assessment as number of students, instructional time, limited training, measurement, and preparation time. These concerns became survey question responses with regard to the organizational aspects of assessment.

When considering the summarizing practices in music assessment, many authors (Brophy, 2000; Fiese & Fiese, 2001; Harrison, 1983; Herrold, 1991; Lavender, 2000; Miller, 2005) identified academic factors that are commonly a part of elementary general music assessment practices. As outlined in the literature, singing, rhythm, melody, form, harmony, instruments, listening, movement, and appreciation are the academic areas that are commonly areas of focus in elementary general music. Therefore, questions on the survey tool will relate to these areas when inquiring about areas of academic assessment.

Additionally, nonacademic factors were also commonly noted as areas relating to summarizing practices in assessment (Allen, 2005; Allen & Lambating, 2001; Barnes, 1985; Brookhart, 1991; Brookhart, 1993; Cizek et al., 1995; Cross & Frary, 1996; Guskey, 1996; Randall & Engelhard, 2010; Wiggins, 1996). Therefore, questions on the survey tool will specifically address participation, attitude, effort, behavior, and punctuality as potential nonacademic factors utilized by Iowa music educators.

In addition to the listed academic and nonacademic factors, several authors (Bouton, 2001; Brophy, 2000; Lavender, 2000) also recommended categorizing factors for clearer organization. The various categories specifically named in the literature also contributed options for participants in a multiple-choice format.

Numerous authors also commonly discussed report cards, conferences, and grades as the most common sharing and reporting practices in elementary music assessment (Allen, 2005; Asmus, 1999; Bailey & McTighe, 1996; Brophy, 2000; Carr & Harris, 2001; Chase, 1999; Cross & Frary, 1999; Fautley, 2010; Guskey, 1996; Harrison, 1983; Lake & Kafka, 1996; MENC, 1996b; Nutter, 1999; Olson, 1995; Power & Chandler, 1998; Shuster et al., 1996; Stiggins, 2005; Tuley, 1985; Tyack & Tobin, 1994). Therefore, questions on the survey tool relating to sharing and reporting practices were oriented toward these three areas – report cards, conferences, and grades.

In addition to the questions relating to Stiggins' (2005) four categories, an additional section of the final survey instrument grew out of the section of the literature review about teachers' motivations, purpose, or reasons for assessment practices. As reflected in the literature review, numerous authors (Barkley, 2006; Brophy, 2000; Brummett, 1993; Farrell, 1997; Fiese & Fiese, 2001; McClung, 1996; Monroe, 1995; Niebur, 1997; Nightingale-Abell, 1993; Schultz, 2002; Talley, 2005) found a variety of rationales relating to assessment in elementary music classes. The ideas shared in the literature provided rich responses to use as multiple-choices on the final survey instrument.

Overall, the literature review served as a guide to designing the 30 questions and the multiple choices on the survey instrument. The close relationship between the final survey questions and the literature lessens the chances of gathering unrelated, superfluous, or meaningless data.

Other surveys. The survey instruments within the related studies included in the literature review provided rich examples of potential survey questions as well as question format possibilities. The formatting from Tally (2005), Barkley, (2006), Cross and Frary, (1996), and Hanzlik, (2001), were examples of clear instructions and concise responses for participants. Specific sources for all content, format, and responses of the questions within the final survey instrument are summarized in Table 2.

Research expert. An additional step in preparing the final survey instrument was consulting with a descriptive research expert (Dr. Barry Wilson, Personal communication, November 28, 2011). Upon initial review, suggestions were made to expand many of the multiple choices and closed response questions to include options for other open responses and opportunities any additional comments to be made by respondents. This may provide rich, direct quotes to include with statistical analyses of the quantities of each option. Also suggested was the reduction of the total number of questions, specifically reducing redundancy from similar questions. The final suggestion was to move the demographic questions to the end of the instrument. A brief, concise explanation of the reason and rationale for the study was notably missing and further recommended as a necessity within the letter to potential participants.

Description of Survey Instrument

A general collection of assessment practices is the overall goal with regard to the research problems stated in this research document. By using the final survey instrument of 30 total questions, this study will aim to analyze the means by which elementary music teachers in Iowa assess their students in the classroom based on Stiggins' (2005) four

categories of assessment: gathering, organizing, summarizing and sharing. Finally, teachers' opinions regarding assessment or motivations for assessing will be targeted, as well as general demographic information of the respondents.

The final survey instrument for this descriptive research study was derived, overall, based on the pilot study, information from the literature review, samples of other survey instruments, and consultation with an expert. The final instrument was also reviewed by a music education colleague, as suggested by Dillman (2000). The colleague consulted on the necessity of each question, checked for clarity of categories and word choice, analyzed the appropriateness of answer choices, and analyzed the potential "impression" (p. 141) of the survey. The revisions that came from these suggestions included eliminating one unclear question, and clarifying the definition of assessment at the onset of the entire instrument.

Reducing or avoiding measurement error was the ultimate goal of the final instrument. Firstly, maximum quality in content of responses with minimum time requirement on the part of the participants was the ultimate goal. Music teachers with already busy schedules will not have time for extensive responses, and a first impression of too much needed information could trigger nonresponsive attitudes. Secondly, minimizing errors in responses was the goal of using a balance of closed- and open-ended responses. The response modes for each question vary according to the nature of each question. Some questions have been designed to elicit responses regarding attitude toward assessment and any conditions or factors that might influence assessment practices. Other questions have been designed to discover categories within certain

settings. Further questions have been designed specifically to gather numbers of occurrences or instances of certain assessment behaviors.

Table 2. *Sources of Survey Instrument Questions.*

Sections of Survey	Survey Question #	Formatting Source(s)	Multiple Choices / Responses Source(s)
Assessment Practices: Gathering	1		Brophy (2000); Fiese & Fiese (2001); Harrison (1983); Herrold (1991); Lavender (2000); Miller (2005)
Assessment Practices: Gathering	2	Barkley (2006)	Pilot instrument; Fiese & Fiese (2001); Herrold (1991); Miller (2005); Wilson, Personal Communication (2011)
Assessment Practices: Gathering	3	Barkley (2006)	Airasian (1994); Berman (2008); Foley (2001); Hill (2008); Salmon-Cox (1981); Stiggins (2005); Taylor (2003)
Assessment Practices: Gathering	4	Talley (2005)	Talley (2005)
Assessment Practices: Gathering	5	Talley (2005)	Shuler (1996a), Colwell (1969), Boyle & Radocy (1987)
Assessment Practices: Organizing	1	Barkley (2006)	Airasian (2000); Brophy (2003); Carter (1984); Colwell (1974); Hamann (2001); Dorr-Bremme & Herman (1986); Lehman (1968); Marzano (2000); Russell & Austin (2010);
Assessment Practices: Organizing	2		Airasian (2000); Brophy (2003); Carter (1984); Colwell (1974); Dorr-Bremme & Herman (1986); Lehman (1968); Marzano (2000); Russell & Austin (2010); Walker (1998)

(Table Continues)

Sections of Survey	Survey Question #	Formatting Source(s)	Multiple Choices / Responses Source(s)
Assessment Practices: Summarizing	1	Barkley (2006)	Brophy (2000); Fiese & Fiese (2001); Harrison (1983); Herrold (1991); Lavender (2000); Miller (2005)
Assessment Practices: Summarizing	2	Barkley (2006)	Allen (2005); Allan & Lambating (2001); Barnes (1985); Brookhart (1991); Brookhart, (1993); Cizek et al., (1995); Cross & Frary (1996); Guskey (1996); Randall & Engelhard (2010); Wiggins (1996)
Assessment Practices: Reporting	1	Barkley (2006)	
Assessment Practices: Reporting	2	Barkley (2006)	
Assessment Practices: Reporting	3		Allen (2005); Asmus (1999); Bailey & McTighe (1996); Brophy (2000); Carr & Harris (2001); Chase (1999); Cross & Frary, (1999); Fautley (2010); Guskey (1996); Harrison (1983); Lake & Kafka (1996); MENC (1996b); Nutter (1999); Olson (1995); Power & Chandler (1998); Shuster et al. (1996); Stiggins (2005); Tuley (1985); Tyack & Tobin (1994)
Assessment Practices: Reporting	5		Pilot instrument

(Table Continues)

Sections of Survey	Survey Question #	Formatting Source(s)	Multiple Choices / Responses Source(s)
Assessment Practices: Reporting	6		Pilot instrument
Teacher Opinions	1		Barkley (2006); Brophy (2000); Brummett, (1993); Farrell (1997); McClung (1996); Monroe (1995); Niebur (1997); Nightingale-Abell (1993); Talley (2005)
Teacher Opinions	2		Barkley (2006); Brophy (2000); Brummett, (1993); Farrell (1997); Fiese & Fiese (2001); McClung (1996); Monroe (1995); Niebur (1997); Nightingale-Abell (1993); Schultz, (2002); Talley (2005)
Teacher Opinions	3	Pilot instrument	Pilot instrument
Teacher Opinions	4	Pilot instrument	Pilot instrument
Demographics and Responsibilities	1 - 11	Pilot instrument	Pilot instrument

Once the questions were formatted electronically, the completed survey was placed on the Internet via an online survey program (Survey Monkey) and also offered to all participants in hard copy form upon request. The complete survey instrument is comprised of 30 questions is attached as Appendix F and will be referred to as the

Elementary General Music Survey on Assessment in Iowa. Specific sources for all of the questions, responses and formatting in the instrument are summarized in Table 2.

To encourage participation and completion of the survey and reduce nonresponse error, the survey respondents were offered a free summary of the research data and final results. Once finalized, an electronic version of the results was e-mailed to those participants, based on their names and e-mail information submitted at the close of the survey. In addition, a noteworthy music store, West Music Company, donated a fifty-dollar gift card to West Music Company. All of the participants who completed the survey were entered into a random drawing for the prize.

Procedure

The researcher obtained permission from the University of Northern Iowa Institutional Review Board to conduct the research. Following approval, a pre-survey communication (Appendix B) was sent to the elementary general music teachers who teach at the 315 schools in the sample. This communication alerted them to the upcoming invitation to formally participate in the research study. One week following, a cover letter (see Appendix C) was sent to the 315 sample schools with a link to the online survey tool. This cover letter was also sent electronically via E-mail. Each respondent was asked to complete the survey online within two weeks.

Each respondent was also offered the optional opportunity to submit a copy of any parent communication tools used in his or her assessment practices, either in hard copy form, sent in an enclosed envelope, or in electronic form, sent as an email attachment. Out of the 211 respondents, no such forms were shared.

The respondents were given two weeks to complete and return or submit the survey. As the two-week period was coming to a close, a reminder letter (Appendix D) was sent to those music teachers who had not yet responded, reminding them of the upcoming deadline. After the deadline had passed, a fourth and final correspondence (see Appendix E) was sent electronically to the remaining teachers in the population who had not yet responded, asking them to complete and submit the survey as soon as possible.

Analysis

The data was collected through the use of the online Survey Monkey tool and was imported into SPSS for analysis. The data from participants who started the survey and did not complete the entire form was deleted.

Descriptive statistics were utilized to describe frequencies of occurrences of many assessment practices as well as sources and categories within assessments, specifically mining the data for as much potential analysis as possible. The statistic that was used for seeking patterns among demographic groups was crosstabulation, or the Chi Square test. The purpose for utilizing such statistical analysis related to the research questions of this study, not only seeking answers but also potential links among variables. The results of data analysis are found in Chapter 4.

Qualitative Study

The second part of this research study was qualitative in nature. Whereas the quantitative data gathered through surveys provided a large amount of information over

many areas and questions from numerous participants, the qualitative questions provided an opportunity to gather more detailed data from a smaller group of participants.

Interviews

Group interviews, or focus groups (Denzin & Lincoln, 2003), provide opportunity for further questions or details relating to the four research areas of this study, namely gathering, organizing, summarizing, and reporting (Stiggins, 2005). This study had two focus groups, one with elementary general music teachers from small school districts and one with elementary general music teachers from a large school district. The overall aim of these groups was to have a positive discussion of what is really happening in the elementary music classrooms with regard to assessment, with regard to the three research questions. The focus groups in this research study provided an opportunity to discover "perspective, experience, and language" (Boeije, 2009, p. 62). The participants' knowledge, reflections, and ideas provided insight to the research questions that this study was designed to explore.

Assessment practices in elementary music was the specific topic discussed in each focus group. The conversation was started by the researcher asking for characteristics of assessment. The goal of the focus group setting was that the participants would be drawn to discussions that evolved and progressed to related issues, challenges, motivators, or benefits. The conversations were rich and informative. The list of the potential protocol can be found in Appendix G. The complete transcripts of each focus group are found in Appendix O and P.

Criteria for interview sample selection. For the two focus groups used in this study, the selection of participants was conveniently sampled from the music teachers in nearby school districts. The two groups had diverse membership in terms of individual demographics, one group included the music teachers from a nearby urban school district with eleven elementary school buildings. Nine of the eleven schools are on Iowa's *Schools In Need Of Assistance* (SINA) current list, with many of the school having been on the assistance list for several years (Iowa Dept. of Education, 2012). The other focus group was made up of music teachers from nearby small, rural schools. This group of individual teachers were the only elementary music teacher in his or her district, and one participant served two buildings within that small district.

Data. The goal of these focus group discussions was twofold. The first goal was to gather any additional data related to the survey questions. The second goal was to open up the general topic of assessment to allow the two respondents to freely contribute any information they feel is relevant or important. The initial engagement of the questions led to further exploration to deeper ideas or perceptions from both groups. All of the focus group members contributed freely and offered extended details beyond the answers on the survey instrument.

Both focus groups were video recorded. After each group discussion, the recordings were transcribed within one week. Following transcription, each group member was sent a copy of the transcription for any comments or corrections. No recommended changes or concerns were shared regarding the content, so no follow-up phone calls were necessary.

Following transcription, the overall goal of analysis was to reduce the bulk of information and seek overlying patterns (Miles & Huberman, 1994). The idea of organizing the data was multifaceted: within each group, and across both groups (rural and urban). There were both commonalities and distinct differences within and across groups.

Initially, the data was reviewed and all patterns or themes were noted. Clustering of commonalities occurred with very few outliers identified. The counting of like ideas also occurred. Large instances of certain comments indicated verification of and supplemented the survey data. These comparisons or contrasts may provide further insight with regard to the research questions as well as the survey data. Summaries of these chains of evidence are summarized in Chapter 4.

Optional Review of Documents

The second part of the qualitative portion of this study was going to include an optional offering from participants to submit any documents relating to assessment practices, specifically reporting assessment data or communicating assessment data to parents. The review of the literature (Chapter 2) indicated various methods and means of summarizing and sharing assessment information and data with parents and other interested parties. The process of document review for this study was planned to aid in identifying themes, or coding, in music assessment documents used in Iowa (Denzin & Lincoln, 2003).

A document guide was developed to focus the review (See Appendix I). This guide, with the above-mentioned attributes as categories, was developed as a tool to

examine with the following specific areas of content: (a) data gathered, unit of study, content area or concepts, (b) organization of form/tool, (c) reporting of nonacademic criterion, and affects the overall/final grade, (d) the types of measurement scales used for students' ratings or grades, (e) comments, in general, written by the teacher for the parents, (f) National Standards for Music Education, and (g) any other information gleaned from the documents. These categorical codes were a "start list" (Miles & Huberman, 1994, p. 58) of potential codes. The goal of this part of the investigation is to document a "literal version of what is there" (Mason, 2002, p. 149).

Unfortunately, out of the 211 respondents to the survey no additional documents were submitted. One respondent did email a snapshot of the format of parent communication. This particular district only utilized a small, open box for limited teacher comments on both content and behavior. With no response with additional documentation, the document guide was not used and this part of the research project was eliminated.

Conclusion

The research design for this inquiry was a mixed methodology approach that used a survey and group teacher interviews as the sources of data. The overall goal of this research was to thoroughly describe what was occurring in elementary general music assessment in Iowa. This multifaceted investigation provided rich details, with much scope and breadth. The next chapter, Chapter 4, presents the results of the data collection and analysis presented in this chapter.

CHAPTER 4

RESULTS

The purpose of this study was to research the current practices in assessment of elementary general music teachers in Iowa. Stiggins (2005) summarized classroom assessment into four categories: gathering evidence, organizing or storing evidence, summarizing evidence, and sharing or reporting evidence. There are three kinds of data that will be shared in this chapter. Firstly, all survey answers from all respondents will be shared – both describing the demographics of the respondents and the answers to each question. Secondly, differences in data based on demographical categories of respondents will be described. Thirdly, data from focus groups of both small and large school districts. Throughout each section, the three research questions will also be addressed.

Quantitative Study: Survey

The sample of 315 elementary music teachers consisted of 35 randomly selected elementary general music teachers from each of the nine Area Educational Agencies (AEAs) throughout the state. A request to complete a survey was sent to the 315 elementary music teachers in the sample. The survey, *Elementary General Music Survey on Assessment in Iowa*, can be found at Appendix C.

When the request for completion was sent to sample 315 recipients, one respondent withdrew from the study with a note explaining that she was no longer taught elementary general music. Another recipient responded that her teaching assignment was in a performing arts academy, and the assessment questions within the survey did not

apply to her job assignment or daily tasks. One other survey was returned with an invalid email address. A paper copy was then sent to this recipient, and that, too, was returned with an insufficient address.

A total of 211 of the 312 surveys were completed and returned to the researcher either online or in the regular mail. This resulted in a 67.63% response rate. Occasionally, a few respondents would skip a question on the survey. The questions that were skipped by some respondents were noted in with the data summaries within this chapter. Overall, the responses from these 211 surveys were utilized as data for this study to provide an understanding of current assessment practices in elementary general music in Iowa with relation to the research questions.

Demographics

The survey instrument included questions regarding the demographics of each respondent. Specific information requested from each participant included school district size, location, and characteristics, years of teaching experience, highest level of education, percentage of full-time equivalency, number of buildings served, grades taught, average number of students taught per week, and average class size. Tables 3 through 12 below are illustrations of the information regarding these demographic areas.

The survey was sent to elementary general music teachers in Iowa of varying sizes. Of the 211 total teachers who completed the survey, 191 respondents self-reported their school district size. The Iowa High School Music Association (2012) divides schools into classifications, depending on the number of students (1A through 4A). Class 1A school districts are those schools with approximately 149 or fewer students in grades

9 to 11 of the previous school year. Class 2A schools are defined as having approximately 150 to 274. Class 3A schools are defined as 275 to 599 students. Class 4A schools have over 600 students. Although the Class A classification was eliminated as a fifth category by the Iowa High School Music Association several years ago, Class A is defined by the Iowa High School Athletic Association. The Iowa High School Athletic Association categorizes Class A schools as the smallest schools in the state ranging from the very smallest enrollments to those schools that are classified as 1A in music (IHSAA, 2013). These specific ranges for school district size were not defined to the instrument, but rather self-reported by each respondent.

Table 3. *Demographic Descriptors of Respondents: School Size and Years of Teaching.* (N = 191, 202)

	Respondents	
Variable	#	%
School District Size		
A	15	7.9
1A	38	19.9
2A	32	16.8
3A	40	20.9
4A	66	34.5
Years of Teaching Experience		
0 to 4	49	24.3
5 to 9	35	17.3
10 to 14	29	14.3
15 to 19	24	11.9
20 to 24	25	12.4
25 to 29	21	10.4
30 to 34	13	6.4
35 to 39	6	3.0
40 to 44	0	0

In this study, 191 respondents indicated school district size, with 15 music teachers reported teaching in a size A school district, 38 taught in a size 1A school district, 32 taught in a 2A school district, 40 taught in a size 3A district, and 66 taught in a 4A school district, as reported Table 3.

The respondents also indicated the total number of years that each had been teaching elementary general music. The years of teaching experience ranged from first year teachers (four) to 38 years (one) of teaching experience, as also indicated in Table 3. The average years of teaching experience was approximately 13.5 years. Nine respondents skipped this question.

Table 4. *Representation of Respondents by Area Educational Agency Region.* (N = 211)

Region	Respondents		
	#	% of 35/AEA	% of Total
AEA 1 - Keystone	26	74.29	12.32
AEA 267	33	94.43	15.64
AEA 8 - Prairie Lakes	17	48.57	8.06
AEA 9 - Mississippi Bend	19	54.29	9.00
AEA 10 - Grant Wood	30	85.71	14.22
AEA 11 - Heartland	31	88.57	14.69
AEA 12 - Northwest	19	54.29	9.00
AEA 13 - Green Hills	14	40.00	6.64
AEA 15 - Great Prairie	22	62.86	10.43

In addition to the survey respondents representing all school sizes in Iowa, the 211 respondents also represented all nine of the Area Educational Agencies (AEA) throughout the state. Thirty-five surveys were distributed within each AEA region with a

range of 14 to 33 respondents from the nine. Table 4 displays all of the respondent numbers and percentages, as well as the percent of each AEA region to the total number of respondents throughout the state. A map of the nine regions is also found in Appendix J.

Table 5. *Demographic Descriptors of Respondents.* (N = 199)

	Respondents	
Variable	#	%
Degree		
Bachelor's	121	60.8
Master's	77	38.7
Doctorate	1	0.5
F.T.E.		
Full-time	133	66.8
Part-time (< 100%)	66	33.2
30 to 44 %	4	6.0
45 to 59 %	30	45.0
60 to 74 %	3	4.5
75 to 89%	24	36.5
90%	5	7.6
Buildings		
One	109	54.8
Two	66	33.2
Three	18	9.0
Four	6	3.0

NOTE: On this chart, "F.T.E." indicates Full Time Equivalency.

Many of the demographic questions were utilized to determine the specific professional teaching assignment situations for each respondent. Out of the respondents that answered this question (N = 199), 121 indicated they had a Bachelor's degree, 77

indicated they had a Master's degree, and 1 respondent indicated that they had a Doctoral degree. Of the 199 respondents who answered this question, 133 indicated they worked full time, 66 indicated they worked part-time, and 12 did not indicate. Of those 66 part-time respondents, full time equivalency (F.T.E) ranged from 30% to 90%. The question regarding number of buildings served was open ended, and all respondents indicated serving in 1, 2, 3, or 4 buildings. The responses varied between full- and part-time respondents. These demographic descriptors of the respondents are summarized in Table 5.

The survey was specifically sent to teachers of general music in elementary buildings throughout Iowa. Not all of the respondents, however, teach music to students at all grade levels. The number of respondents and percent of overall respondents who reported teaching at each of the designated elementary grade levels is summarized in Table 6. Over 90% of the respondents indicated teaching Kindergarten through fourth grade music classes. Only 26.3% indicated teaching Pre-Kindergarten; 85% indicated teaching fifth grade; and 42.4% indicated teaching sixth grade.

Also of interest in this study was the total amount of students that each respondent taught, as well as the average number of students within each music class that each respondent taught. Also in Table 6, over half of the respondents taught between 300 and 500 students in any given week of teaching elementary general music with 33, or 16.6%, teaching over 500 students. Only 3 respondents indicated teaching 100 or less students per week. Also indicated in Table 6, a majority (61.7%) of the responding teachers had approximately 21 to 25 students in each elementary general music class.

Table 6. *Teaching Variables of Respondents.* (N = 198. 199, 201)

Variable	Respondents #	%
Grade Levels		
Pre-Kindergarten	52	26.3
Kindergarten	184	92.9
First Grade	181	91.4
Second Grade	186	93.9
Third Grade	180	90.9
Fourth Grade	181	91.4
Fifth Grade	169	85.4
Sixth Grade	84	42.4
Number of Students		
100 or less	3	1.5
101 to 200	20	10.1
201 to 300	36	18.1
301 to 400	53	26.6
401 to 500	54	27.1
Over 500	33	16.6
Class Size		
11 to 15 students	1	0.5
16 to 20 students	29	14.4
21 to 25 students	124	61.7
26 to 30 students	46	22.9
31 to 35 students	1	0.5
36 to 40 students	0	0.0

Full-Time equivalency and number of students. Many of the responses in the Tables above may be crosstabulated or referenced to provide additional insight. Combining number of student responses with responses of full- and part-time responses further describe teaching loads of respondents. As indicated in Table 7, a majority of the teachers who taught fewer students were also assigned a part-time teaching assignment. All of the teachers who saw 100 or fewer students were only assigned a part-time

teaching assignment. Approximately 83% of the 23 respondents who indicated teaching less than 200 students total were teaching part-time assignments. Furthermore, all but one, or 97% total, of the respondents who indicated teaching more than 500 students were full-time teachers.

Table 7. *Students Taught Per Week by Teaching Load.* (N = 199)

Number of Students	Respondents' Load		
	Full	Part	Total
100 or less			
Count	0	3	3
Percentage %	0	100	100
101 to 200			
Count	4	16	20
Percentage %	20	80	100
201 to 300			
Count	17	19	36
Percentage %	47	53	100
301 to 400			
Count	34	19	53
Percentage %	64	36	100
401 to 500			
Count	46	8	54
Percentage %	85	15	100
Over 500			
Count	32	1	33
Percentage %	97	3	100
TOTAL	133	66	199

Research Questions

The survey was designed to collect data specifically addressing the research questions.

1. What are the features of the classroom assessment practices currently implemented to:
    a. gather evidence (assessment data) of student learning in elementary general music in Iowa?
    b. organize student learning data in elementary general music in Iowa?
    c. summarize assessment data in elementary general music in Iowa?
    d. report assessment data in elementary general music in Iowa?
2. What demographics, or teacher characteristics, influence or relate to classrooms assessment practices of gathering, organizing, summarizing, and reporting assessment data in elementary general music in Iowa?
3. What are teacher beliefs or opinions related to assessment in elementary general music in Iowa?

Data from the respondents' surveys was used to answer these questions.

Question 1A: Practices of gathering assessment data. The first research question was addressed by several of the questions on the survey instrument. Descriptive analytic techniques were utilized to express the data in meaningful ways. The frequencies of the assessment practices indicated by participants are found in Table 11. As shown in Table 8, over 96 percent of the 211 respondents indicated frequently utilizing observations in the elementary general music classroom to gather evidence of student learning. Portfolios were used least often, with a majority (77.7%) of the respondents indicated having never used them as a tool to gather evidence of student learning.

Table 8. *Frequencies of Respondents' Assessment Tools.* (N = 211)

Tool	Reported Frequency N	S	O	F	M	Mdn	SD
Examination							
Count	28	68	97	18	2.50	3	0.82
Percentage %	13.3	32.2	46.0	8.5			
Projects							
Count	32	58	93	28	2.55	3	0.90
Percentage %	15.2	27.5	44.1	13.5			
Assignment / Homework							
Count	69	62	59	21	2.15	2	0.99
Percentage %	32.7	29.4	28.0	9.9			
Observation							
Count	3	0	4	204	3.94	4	0.38
Percentage %	1.4	0.0	1.9	96.7			
Audio/Visual Recordings							
Count	36	39	71	65	2.78	3	1.06
Percentage %	17.1	18.5	33.7	30.8			
Concert Performance							
Count	21	51	76	63	2.86	3	0.96
Percentage %	9.9	24.2	36.0	29.9			
Portfolio							
Count	164	18	22	7	1.39	1	0.80
Percentage %	77.7	8.5	10.4	3.3			
Rubric							
Count	43	55	76	37	2.51	3	1.00
Percentage %	20.4	26.1	36.0	17.5			
Other			1	8			

NOTE: from the survey instrument, this chart summarizes frequency abbreviations for N as never, S as seldom, O as occasionally, and F as frequently. When calculating mean (M), median (Mdn) and standard deviation, N = 1, S = 2, O = 3, and F = 4.

When considering specific music objectives, the respondents also indicated which practices they commonly used for each objective. Similar to the results in Table 8, observations were reported as the most prevalent assessment tools by the respondents in this sample. The respondents marked as many different types of assessment tools as they

use with each concept. Therefore, many respondents indicated multiple tools used for each objective. The data in Table 9 indicate that some concepts are being assessed more than others. Steady beat, for example, is assessed often while composers are not.

Table 9. *Assessment Tool Use by Music Objective.* (N = 210)

	Assessment Tool								
Objectives	E	Pr.	A/H	O	A/V	C	Po.	R	NA
Keeping a steady beat	32	29	10	201	47	97	12	46	3
Changing tempo	39	32	32	156	47	49	7	20	37
Rhythm identification	100	58	53	178	38	51	13	42	5
Rhythmic accuracy	71	49	42	198	64	88	6	56	4
Time signature	72	48	48	111	24	38	11	18	30
Melodic contour	64	43	42	142	32	33	12	17	23
Note Names	123	57	85	112	16	32	15	31	16
Melodic accuracy	49	23	11	177	38	89	8	37	10
Major / minor tonality	42	12	16	88	28	20	4	7	90
Instrument families	105	75	70	119	54	11	10	19	8
Pitched/unpitched perc.	46	51	34	145	48	35	2	13	22
Recorder	69	55	57	137	30	76	6	55	41
Composers	45	53	47	41	35	10	5	16	87

NOTE: from the survey instrument, this chart summarizes abbreviations for assessment tools as Examinations (E), Projects (Pr.), Assignments/Homework (A/H), Observations (O), Audio / Visual Recordings (A/V), Concerts (C), Portfolios (Po.), Rubrics (R), and Not Assessed (NA).

Another option for respondents on this question was to indicate whether or not a particular objective was not assessed. Out of the 13 objectives that were specifically described, the 210 respondents (on this question) indicated 374 different times that certain objectives were not assessed, as displayed in Table 9. Ninety of the respondents indicated that major and minor tonalities were not assessed and 87 respondents indicated

not assessing composers. Only 3 respondents reported not assessing steady beat, and 4 and 5 respondents reported not assessing rhythmic identification and rhythm performance accuracy, respectively.

The specific question on the survey instrument that addressed this data also allowed respondents to write in any other comment or related idea. The comments in the "other" category included:

- I also assess Form.
- I use marker boards and have students answer questions on them. I can see if they know the answer.
- Fourth and Fifth graders keep portfolio of worksheets, tests and music that is shared with parents at the end of the year.
- Also assess mallet technique.
- I assess improvising with a rubric.
- I assess simple composition using rubrics.
- I also assess articulation, form, tempo and dynamics.
- I use melodic and rhythmic dictation.
- I use checklists, presentations, and discussions.
- I assess solfege.

The respondents also indicated any use of commercially available tests. Sixteen of the 211 respondents indicated using a commercially available test for assessment purposes in the elementary general music classroom. The specific results are noted in Table 10. As indicated, a majority of the respondents indicated no use of the commercial tests. All of the responses came from 16 total, individual respondents. The *Silver*

Burdett Music Competency Test was reported used most often, even though only reported used by 11 different respondents.

Table 10. *Commercial Test Use of Respondents.* (N = 211)

Test	Not used	Grade						
		K	1	2	3	4	5	6
Musical Aptitude Profile	206	-	1	-	-	1	2	1
Primary Measures of Music Audiation	208	1	1	-	1	-	-	-
Intermediate Measures of Music Audiation	209	-	-	-	-	1	1	-
Music Achievement Test	208	-	-	-	1	1	2	-
Iowa Tests of Music Literacy	208	-	-	-	-	1	1	1
Drake Music Aptitude Tests	210	-	-	-	-	-	-	1
Wing Standardized Test of Music Intelligence	211	-	-	-	-	-	-	-
Measures of Musical Abilities	210	-	-	1	1	-	-	-
Watkins/Farnum Performance Scale	210	-	-	-	-	1	-	-
Farnum Music Tests	211	-	-	-	-	-	-	-
Indiana-Oregon Music Discrimination Tests	211	-	-	-	-	-	-	-
Simons Measurements of Music Listening	210	-	-	1	-	-	-	-
Silver Burdett Music Competency Tests	194	3	3	1	4	2	3	1

The respondents also indicated how many formal assessments they had administered since the onset of the school year (approximately 4 months). The survey question further asked respondents to delineate numbers of formal exams by grade level. If a respondent did not teach in a given grade level, “not applicable” or N/A was to be selected. The responses are found in Table 11.

As the term “formal assessment” was defined at the onset of the survey instrument as the “planned and organized use of measurement tools in the classrooms where student learning is documented,” many respondents indicated that zero formal assessments had occurred in the four months since the start of the school year. A total of 31 respondents indicated they had used more than 16 assessments since the onset of the school year.

When asked about the content of assessments, respondents were asked to specifically identify types of music objectives that are assessed in elementary general music. The choices were singing, rhythm, melody, form, listening, harmony, music appreciation, instruments / timbre, and movement. The teachers were asked to

Table 11. *Frequencies of Formal Assessments by Grade Level.* (N = 210)

		Number of Assessments					
Grade	DNT	0	1 - 4	5 - 8	9 - 12	13 - 16	>16
Kindergarten	26	49	75	41	14	4	1
First	29	33	99	37	9	2	1
Second	24	28	84	33	18	15	8
Third	30	25	67	38	31	12	7
Fourth	29	27	66	45	31	8	4
Fifth	41	26	64	34	26	12	7
Sixth	126	22	29	21	7	2	3

NOTE: respondents indicated the number of formal assessments utilized since the beginning of the school year for each of the grades taught. If they did not teach that grade, this was also reported and is documented in the table as DNT, for “do not teach.”

identify as many that applied. The results of this question are summarized in Table 12. A majority, 89% or above, of the respondents assess singing, rhythm, and form. Another large number of respondents (76.1%) indicated assessing melody and instruments or timbre. The content areas with the lowest responses were harmony (49.8%), listening (58.4%), music appreciation (40.7%), and movement (38.8%).

This question was very similar to the survey data shared in Table 9. The duplication of question was purposeful, this time without any indication of how, or which tool was used, the concepts were assessed, but simply whether or not they were assessed at all. The intention was to see if the respondents reported similar data when asked about content in two different scenarios. In comparison, Harmony (major and minor) and music appreciation (composers) were again reported very low in reported assessment. Again, rhythmic and melodic concepts were predominately reported being assessed by the respondents.

Table 12. *Types of Music Objectives Assessed.* (N = 209)

Objective	Respondents	
	#	%
Singing	208	99.5
Rhythm	207	99.0
Melody	159	76.1
Form	186	89.0
Harmony	104	49.8
Instruments / Timbre	159	76.1
Listening	122	58.4
Music Appreciation	85	40.7
Movement	81	38.8

Question 1B: Practices of organizing assessment data. Many questions on the survey instrument were intended to document respondents' practices of organizing assessment data in the elementary general music classroom. The data summaries in Table 13 reflect teacher agreement or disagreement with general assessment organizational practices.

A majority of respondents indicated that the large numbers of total elementary students taught created limitations in assessment organization efforts, with 143 respondents having indicated "agreed" or "strongly agreed" with this statement and another 23 respondents were "unsure." Out of the 207 respondents that answered these

Table 13. *Respondents' Practices in Organizing Assessment Data* (N = 207)

Organizational Statement	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	StDev
Assessments are limited due to large numbers of students.	9	32	23	87	56	3.71	4	1.14
I received adequate training in assessment tools	65	92	29	16	5	2.05	2	0.99
I have PD opportunities in music education assessment	70	87	21	23	6	2.07	2	1.07
Adequate classroom time with my students to assess	99	89	10	6	3	1.67	2	0.82
I could complete more thorough assessments in music class if I had more preparation time outside of class	14	45	36	74	38	3.36	4	1.20

NOTE: PD = professional development. Statistical analyses were abbreviated as M = mean. Mdn = Median. StDev = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

opinion questions, 157 shared that they did not receive adequate training in assessment tools and, furthermore, another 157 indicated the lack of professional development opportunities to grow in assessment knowledge. When considering the organization of assessment data, 109 respondents declared a lack of preparation time (outside of the classroom) to organize assessment data. Likewise, 112 respondents shared that organizational assessment practices could be more "complete" with additional "preparation time outside the classroom."

Table 14. *Respondents' Reported Tendencies With Organizing Assessment Data.*

Survey Statement	Respondents	
	#	%
Collecting data in multiple sections	163	82.7
Managing data for large numbers of students	143	72.6
Identifying strengths and weaknesses	153	77.7
Analyzing trends by class or grade	108	54.8
Transferring observation to organized system	84	42.6

The next survey question regarding assessment organization asked respondents to indicate which, of a set, of practices regularly occurred in elementary general music. Out of 197 respondents for this particular question, 163 respondents, or 82.7% indicated "collecting data." "Managing data" was reported by 143 respondents, or 72.6%. Using the data to "identify areas of strength and weakness" was indicated by 153 respondents,

or 77.7%. "Analyzing" or "transferring observation notes" were reported by 108 respondents, or 54.8%, and 84 respondents, or 42.6%, respectively. These questions were aimed to seek numbers of elementary general music educators who participate in practices related to organizing assessment data and are summarized in Table 14.

Question 1C: Practices of summarizing assessment data. The survey instrument included many statements that asked respondents for their opinion regarding summarizing factors in elementary general music assessment, Stiggins' third facet of assessment. The statements were created with the intention to determine what data, if any, respondents were using when summarizing for assessment purposes. Specifically, with regard to summarizing assessment data, this study is seeking to determine which academic areas and which nonacademic areas are used when summarizing assessment data. Again, to differentiate between organizing and summarizing, this third step of summarizing specifically targets the academic and nonacademic factors considered in the entire assessment process.

The first question regarding summarization asked respondents which of the major content areas of elementary general music education were important. When considering specific skills or concept areas, a large majority of the respondents agreed that singing, rhythm, melody, and listening were "important factors to consider when assessing elementary music students," as 172, 177, 173, and 166 (out of 198 answering this question), respectively, had reported. When asked about listening, 166 respondents indicated that this was an "important factor." Still a majority, movement, form and instruments were also indicated as "important factors" with 128, 123, and 121

respondents respectively. Music appreciation and harmony still had a majority of respondents indicating "importance," with 106 and 105 respondents indicating "importance." This data is summarized in Table 15.

From Table 15, one can learn from the reported data that, for the most part, the respondents consider the listed content areas important, with each mean being greater than 3. Even though the means are above 3, the standard deviations do vary, with both movement and music appreciation standard deviations (1.00 and 1.02 respectively) indicating larger variation. Another important aspect to learn from the data in Table 15 is

Table 15. *Respondent Opinions Regarding Assessment Summaries.* (N = 198)

Curricular Areas for Consideration of Importance	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	StDev
Singing	2	3	21	111	61	4.14	4	0.74
Rhythm	1	1	19	99	78	4.27	4	0.69
Melody	2	3	20	123	50	4.09	4	0.70
Form	2	18	55	104	19	3.61	4	0.82
Harmony	1	32	58	95	12	3.43	4	0.85
Instruments	2	21	53	88	34	3.66	4	0.92
Listening	2	7	23	99	67	4.13	4	0.82
Movement	2	24	44	76	52	3.77	4	1.00
Music Appreciation	10	16	66	74	32	3.52	4	1.02

NOTE: Statistical analyses were abbreviated as M = mean. Mdn = Median. StDev = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

the large number of respondents that indicated “unsure” as their response. This self-reporting would indicate either the respondent did not understand what was being asked, did not understand the question, or perhaps that they are not assessing the item and did not know how to respond. There are significantly large numbers of respondents who indicated the “unsure” option when responding to whether or not each academic concept was considered important, with most at approximately 10% to 33% (harmony) reporting “unsure.”

Duplication is, again, apparent in this section of the survey data. I was asking, at this point in the survey, if the respondents deemed each curricular area important when summarizing assessment data. This connects directly to previous Tables 9 and 12 where respondents originally indicated frequencies in gathering the data. Table 15 connects to this previous data, having asked the respondents if they use the collected data in summary efforts. In comparison, melodic and rhythmic concepts are again reported as being used in summarization efforts. Once again, music appreciation (composers) and harmony (major / minor) were the least reported concepts in summarization efforts. This is consistent throughout the three related questions.

In addition to indicating what the respondents deemed valuable when summarizing elementary general music assessment data, the data gathered from this question on the survey also indicated what factors respondents deemed unimportant. Although not a large percentage or near a majority, five teachers indicated “singing” not to be an important factor of elementary music assessment. Two teachers indicated “rhythm” not being important. Five teachers indicated “melody” as unimportant. When

considering the teaching of “form,” 20 teachers indicated it was not an important factor. “Harmony” was marked unimportant by 33, the “instruments” category was marked unimportant by 23 respondents, “listening” by nine, “movement” by 26, and “music appreciation” by 26 total respondents.

Moreover, many respondents indicated that they were “unsure” about certain concepts being important to elementary general music assessment. The “unsure” category was chosen 374 times when respondents were asked about curricular concepts being important in music assessment. This continues to indicate a lack of clear understanding among respondents regarding how to summarize assessment data and what should be considered when summarizing assessment data.

The next survey question regarding summarizing assessment data was designed to elicit views from elementary general music teachers about how much nonacademic factors, or behavioral aspects, contributed to the summarizing practices. Behavior was broken into subcategories of participation, attitude, effort, conduct, and punctuality. All of the responses related to this second survey question regarding behavioral important on summarizing assessment data are summarized in Table 16. When asked about each concept and the relative importance, 186 considered “participation” important, 186 considered “attitude” important, 177 considered “attitude” important, 189 considered conduct important, and 55 considered “punctuality” important criteria when assessing elementary general music students. Each concept also had elementary music teachers who indicated no importance when considering assessment. Seven respondents indicated they did not believe the behavioral factors, together, were important in assessment.

When asked about each concept individually, 9 indicated that "participation" was not important, 11 indicated that "attitude" was not important, 5 indicated that conduct was not important, and 55 indicated that "punctuality" was not an important factor when regard to music assessment. Furthermore, 139 total respondents indicated that they were "unsure" about the importance of these concepts, both together and individually, when considering elementary general music assessment.

Table 16. *Opinions on Behavioral Factors of Assessment Summaries.* (N = 201)

Behavioral Factors for Consideration of Importance	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	StDev
Participation is important	3	6	6	95	91	4.31	4	0.80
Attitude is important	3	8	13	111	66	4.14	4	0.82
Effort is important	2	3	7	110	79	4.30	4	0.79
Conduct is important	3	7	13	101	77	4.25	4	0.84
Punctuality is important	21	34	91	45	10	2.94	3	1.00

NOTE: Statistical analyses were abbreviated as M = mean. Mdn = Median. StDev = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

To further delineate respondents' ideas regarding academic vs. nonacademic factors in assessment, I added two more questions – one at the end of the academic factor question and one at the end of the nonacademic factor. At the conclusion of the first question, respondents answered: "if a student tries hard but performs poorly on music objectives, he/she should receive a poor grade." At the conclusion of the second question,

respondents answered: “if a student tries hard but performs poorly on music objectives, he/she should still receive a good grade.” This was a concluding point each time, truly aimed to investigate the extent to which teachers believe the relationship between academic and nonacademic factors should be. The difference in wording specifically related to either emphasizing only academic factors (first question) or including nonacademic factors (second question).

Table 17. *Two Questions on Behavioral Factors of Assessment Summaries.* (N = 201)

Behavioral Factors for Consideration of Importance	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	StDev
“If a student tries hard, but performs poorly on music objectives…” would have a NEGATIVE impact on student’s assessment.	24	47	89	24	14	2.78	3	1.04
“If a student tries hard, but performs poorly on music objectives…” would have a POSITIVE impact on student’s assessment	17	42	55	67	20	3.15	3	1.13

NOTE: Statistical analyses were abbreviated as M = mean. Mdn = Median. StDev = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

A large number of teachers were unsure about including these nonacademic, behavioral factors when summarizing assessment data, where 89 (or 44%) and 55 (or 27%) indicated “unsure” about including nonacademic factors. Seventy-one teachers

(35%) indicated that if student behaviors are positive, that should have a positive impact on a student's assessment even when academic achievement is low. Similarly, almost half (43%) indicated that positive behaviors should have a positive impact on student assessment. These summaries are found in see Table 17.

Question 1D: Practices of reporting assessment data. Many questions on the survey instrument were created to seek information regarding practices of reporting assessment data to parents. These questions were not written with the assumption that (a) reporting took place, or (b) reporting included sharing of actual data. Respondents were asked questions about communication, frequency, specific reporting tools, conferences, report cards, and formatting.

Table 18. *Practices Relating to Reporting Data.* (N = 201)

Assessment Reporting Statement	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	StDev
"I collect enough data to accurately communicate student growth"	12	107	40	35	7	2.60	2	0.96
"I could provide adequate documentation if a parent questioned student grade"	5	72	74	42	8	2.88	3	0.90

NOTE: Statistical analyses were abbreviated as M = mean. Mdn = Median. StDev = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

The first question regarding reporting assessment data to parents asked about quantity of assessment used in reporting and confidence of using that data to justify grading practices. When asked about the sufficiency of the quantity of data collected when reporting to parents, 119 respondents, or 59%, indicated they did not have enough data. Only 42, or 21%, respondents indicated having enough data. Forty respondents, or 20%, indicated they were "unsure." When asked about the level of confidence in justifying a student's grade with supporting data or documentation, 77 respondents, or 38%, indicated they could not provide documentation in support of their grading practices. The "unsure" option was marked by 74 respondents, or 37%. Only 42 respondents marked "agree" and only eight marked "strongly agree" to the statement of providing documentation to support grading practices. These responses are summarized in Table 18.

Respondents were also asked about reporting tools utilized in their elementary general music practices, specifically the types of tools and frequency of use. Respondents were given the choices of "report cards," conferences," and "grades," with which they were to respond on frequency of "never," seldom," "occasionally," or "frequently." The was also an "other" category where each teacher had the option of inputting another specific type of reporting tool or practice.

As shared in Table 19, a large majority of respondents indicated use of report cards with 77 reporting using them "occasionally" and 91 reporting using them "frequently." Only 15 respondents indicated the "frequent" use of conferences and 61 indicated "occasionally" utilizing conferences. A majority, 113 respondents, indicated

"seldom" or "never" using conferences. Although grades or grading would technically fall under a way to summarize assessment, "grades" was included in this question as a means of reporting or communicating progress to parents. When asked about using grades in elementary general music, 61 indicated "never" using grades, 69 indicated a "seldom" use of grades, 44 indicated "occasionally" using grades and only 11 indicated a "frequent" use of grades. Some respondents may have not realized that the "never" response choice indicated not using or not applicable. Not all of the respondents shared a response on this question for each of the categorical options.

Table 19. *Frequency of Using Reporting Tools.* (N=202)

	Reported Frequency						
Tool	N	S	O	F	M	Mdn	StDev
Report Cards (N = 202)							
Count	12	22	77	91	3.22	3	0.87
Percentage %	5.9	10.9	38.1	45.1			
Conferences (N = 189)							
Count	21	92	61	15	2.37	2	0.78
Percentage %	11.1	48.7	32.3	7.9			
Grades (N = 185)							
Count	61	69	44	11	2.02	2	0.90
Percentage %	33.0	37.3	23.8	5.9			
Other (N = 24)							
Count	10	3	3	8	2.38	2	1.35
Percentage %	41.7	12.5	12.5	33.3			

NOTE: from the survey instrument, this chart summarizes N as never, S as seldom, O as occasionally, and F as frequently. When calculating mean (M), median (Mdn) and standard deviation (StDev), N = 1, S = 2, O = 3, and F = 4.

One last option for respondents regarding types and frequency of reporting options was an “other” category, which was selected by 24 respondents. Of these respondents, 16 shared written comments of what was used as this “other” choice. Responses included:

- Four quarters for report cards
- Conferences are open, but parents hardly come to see me
- I email parents (2 comments)
- Concerts (4 comments)
- Rubrics (3 comments)
- S, I and U
- Music Class Points
- Individual student goals
- Phone calls to parents (2 comments)

Although not identical, two similar comments indicated emailing parents, four comments identified concerts as the means of reporting progress to parents, three indicated rubrics, and two noted phone calls to parents.

For those teachers that did indicate using report cards, the following question on the survey instrument asked how student progress was indicated or communicated on report cards. The options were letter grades, numeric rating scale, descriptive rating scale, plus/check/minus, or emerging/competent/mastery. Again, there was also an “other” category for respondents to further explain a different option for specific communication categories on report cards. Summaries of the responses to how elementary general music report cards are categorically reported to parents are found in Table 20.

Table 20. *Reported Categorical Options for Report Card Use.* (N=196)

Options	Respondents #	Respondents %
Letter grades (A,B, C, D, F)	28	14.0
Rating scale (1, 2, 3, 4, 5)	35	17.5
Descriptive (excellent/good)	71	35.5
Plus/Check/Minus	19	9.5
Emerging/Competent/Mastery	20	10.0
Other	23	11.5

When asked about the specific categories utilized in report cards, 28 respondents indicated using “letter grades,” 35 indicated a numerical “rating scale,” 71 indicated a descriptive rating of “excellent and good,” 19 indicated a “plus/check/minus” system, 20 indicated “emerging/competent/mastery” word use, and 23 indicated “other” uses and stated the specifics. Within the “other” category, several respondents (11) indicated a varied choice of a descriptive label for categories of competency, such as: satisfactory/needs-improvement/unsatisfactory, support/sometimes/consistently, acquired/developing/beginning, beginning/developing/secure, growth/competent, needs/meets/exceeds, mastery/progressing-towards, and advance/proficient/basic/below. A few respondents (3) indicated the use of rubrics. Finally, eight of the respondents indicated “comments” or “personal comments” as their means of communication on report cards.

Further questions regarding report cards were found on the survey instrument and respondents were asked to indicate whether or not their school district provided a standardized report card for use. Additional follow up questions provided the opportunity for respondents to indicate if one standard form was utilized throughout each school year or if the form changed with each reporting period. Out of 205 respondents, 124, or 60.5% indicated teaching in a district that mandated a common, district-wide form for all elementary general music teachers. Sixty-nine, or 33.7% of the teachers indicated creating their own, individual form for use as a report card in elementary general music. These answers are summarized in Table 21. Again, this number of respondents is consistent with the number that reported using report cards in previous survey questions.

Table 21. *Report of District-Wide Mandated Report Cards*. (N = 205)

	Respondents	
Options	#	%
YES – district mandated common form	128	62.4
NO – individualized form created	69	33.7
I do not use any form of report card	8	3.9

Further questions regarding the format of report cards were answered on the survey instrument. Respondents were asked to indicate if the report card format they used was the same or different throughout each reporting period of the year, and if it was

the same or different for each grade level taught. Changes in the card could be referenced to change in curricular units throughout a school year, or to age-appropriate

Table 22. *Reported Changes in Report Card by Time or Grade*. (N = 192)

Change Options	Respondents	
	#	%
Changed throughout the school year	36	18.8
No change through the school year	156	81.3
Varied for each grade level	123	64.1
Not varied for each grade level	69	35.9

expectations or wording of concepts. Out of 192 respondents, a large majority – 156, or 81.3% – indicated that the report card format they used remained "exactly the same for every conference reporting period." A majority – 123 or 64.1% – indicated using a varied form "for each grade level" taught. These responses are summarized in Table 22.

The elementary general music teachers responding to the survey were further asked to comment on exactly what changes occurred (if any) throughout the year and with grade levels. The comments – condensed and counted – were:

- music units, concepts, elements since last reporting period (13 comments)
- mark with N/A any concepts we haven't covered (6 comments)
- whether there has been a concert or not
- I expect better behavior as the year progresses
- music concepts assessed at semester; behavior assessed at quarter.

The last question on the survey instrument with regard to reporting assessment data to parents was written to seek information about parent-teacher conferences. Information was sought about if the conferences exist, and, if so, what role the music teacher was to play during the conference time. Of the 198 teachers that responded to this survey question, 184 of these respondents reported having to be "present at school during conferences." Many of these, 93, indicated having to be in "a certain location at the school during conferences." When asked about whether or not parents actually visit them during conferences, 119, or approximately 60.1%, of the respondents indicated positively that parents did seek to find them and discuss student progress. A large majority (151) indicated that the required workplace during conferences was in their music classroom or office. Only five respondents indicated not being required to be present at conferences. These responses are summarized in Table 23.

Table 23. *Reported Requirements During Conferences.* (N = 198)

	Respondents	
Change Options	#	%
Required to be present during conferences	184	92.9
Specific location required during conferences	93	47.0
Visited by parents during conferences	119	60.1
Allowed to work in office/classroom	151	76.3
Not required to be present at conferences	5	2.5

The teachers that indicated that they were required to be in a "certain location at the school" during conferences were also offered the opportunity within the survey instrument to report the exact location. Of the 93 respondents who indicated a required location during conferences, 87 specified the exact location with 16 reporting the school gymnasium, 24 reporting the school library, and 47 reporting the elementary music room.

Question 2: Demographic comparisons with responses. To answer the second research question, variables were analyzed to determine any significant differences among respondents of varying demographic categories. Data was coded and entered into a computer statistical program, SPSS (2013), for analysis. Statistical analyses were used to compare and discover relationships or differences specifically between demographic characteristics of the respondents and the multiple responses to research Question 1 (practices of gathering, organizing, summarizing, and reporting assessment data). The demographic categories that were investigated were school size, number of students, degree, number of buildings, and years of teaching. The survey questions used for analysis were regarding teachers' practices for gathering data, organizing data, and summarizing data. The specific data used for gathering, organizing, and summarizing assessment data was taken directly from the survey questions, specifically survey question number 2 in section 2, regarding collecting (gathering), managing (organizing), and analyzing (summarizing) data.

School size. In a comparison of respondents from varying sizes of school districts, further significant differences were sought. I specifically wanted to investigate whether or not the respondents' school district size made an impact on their answers and

if any trends or patterns occurred in answers depending on school district size. For the sake of statistical analysis, "small school districts" was assigned to respondents with size A, 1A, and 2A school. Respondents were considered from "large school districts" at the 3A and 4A size. The total number of respondents teaching in small schools was 85, and the total number of large school respondents was 106.

When answering the first research question, *What are the practices currently implemented to gather, organize, summarize, and report assessment data in elementary general music in Iowa?*, the researcher conducted Chi-Square statistic between the size of school district and the respondents' indications of gathering, organizing, and summarizing assessment data. Because it could be argued that Pearson Product was not a viable calculation to use with categorical data and small numbers of categories within each set of data, I felt that the Chi-Square statistic could tell the same story. With the data being close to interval data, I also wanted to check assumptions and ran both non-parametric and parametric (Pearson) statistics. I found the same results to be true. But, for the purpose of this research, I felt like the Chi-Square statistical test would describe similarities and differences among groups very well.

When calculating the Chi-Square statistic, I was seeking how likely the change in respondents' answers regarding gathering, organizing and summarizing data was due to chance or if there was a statistical significance in the change of respondents' answers based on school size. None of the assessment practices were found statistically significant with relation to school size, as shared in Table 24.

Table 24. *Crosstabs of Demographic Variables and Assessment Practices.*

Variables	Chi^2
School Size	
Gathering Assessment Data	0.223
Organizing Assessment Data	0.196
Summarizing Assessment Data	0.007
Number of Students	
Gathering Assessment Data	7.560
Organizing Assessment Data	2.250
Summarizing Assessment Data	2.780
Number of Buildings	
Gathering Assessment Data	2.465
Organizing Assessment Data	2.693
Summarizing Assessment Data	0.007
Years of Teaching Experience	
Gathering Assessment Data	4.837
Organizing Assessment Data	12.862**
Summarizing Assessment Data	22.665**

NOTE: * $p < 0.05$ ** $p < 0.01$

Number of students. Further statistical analyses were conducted in comparing respondents' answers considering the numbers of students taught. The data indicated possible significant differences between respondents with lesser total number of students and respondents with higher total number of students. For the sake of statistical analysis and reducing the degrees of freedom and minimum responses of each level of numbers of students, the lowest two levels of numbers of students were combined. Thus, "less than 100" and "101 to 200 students" were combined to "under 200" total students.

When answering the first research question, *What are the practices currently implemented to gather, organize, summarize, and report assessment data in elementary*

general music in Iowa?, I calculated the Chi-Square statistic between number of students and each assessment practice. In calculating the Chi-Square statistic, I was seeking how likely the change in respondents' answers regarding gathering, organizing and summarizing data was due to chance or if there was a statistical significance in the change of respondents' answers based on total number of students taught. Again, none of the assessment practices were found statistically significant with relation to total number of students taught, as shared in Table 24.

Number of buildings. A comparison of respondent answers was also compared with respect to the number of buildings each responded served or in which they taught. On the survey, respondents indicated the total number of buildings they travel to for their full- or part-time equivalency job. The data was used to seek potential significant differences between respondents who remain in one building each day and those respondents who travel between two or more buildings. The total number of respondents teaching in one building was 109 (54.8%), and the total number of respondents traveling between multiple buildings was 90 (45.2%).

When answering the first research question, *What are the practices currently implemented to gather, organize, summarize, and report assessment data in elementary general music in Iowa?*, I again calculated the Chi-Square statistic, seeking how likely the change in respondents' answers regarding gathering, organizing and summarizing data was due to chance or if there was a statistical significance in the change of respondents' answers based on total number buildings served. Again, none of the

assessment practices were found statistically significant with relation to total number of buildings served, again shared in Table 24.

Teaching experience. Significant differences were also sought when variables were analyzed by years of teaching experience. The data revealed possible significant differences between less experienced respondents and respondents with greater years of teaching experience. The respondents, on the survey, reported their years of teaching in an open ended question. For analysis, the respondents' years of teaching experience were grouped in spans of 0 to 9 years, 10 to 19 years, 20 to 29 years, and 30 or more years of teaching experience. Within these parameters, 84 respondents had taught less than nine years; 53 respondents had taught between 10 to 19 years; 46 respondents had taught between 20 and 29 years; 19 respondents reported having taught over 30 years. These four quantities were then compared to respondents' self-reported practices of gathering, organizing and summarizing assessment data from Research Question number one.

The researcher calculated a Chi-Square statistic between school size and each assessment practice to test the null hypothesis. The results are presented in Table 24. When calculating the Chi-Square statistic, the researcher was seeking how likely the change in respondents' answers regarding gathering, organizing and summarizing data was due to chance or if there was a statistical significance in the change of respondents' answers based on total number buildings served. A Chi-Square test was also performed and statistical significance was found between years of teaching experience and both assessment organization, $\chi^2(4, N = 188) = 12.862$, $p < 0.01$, and summarization, $\chi^2(4, N = 188) = 22.665$, $p < 0.01$.

Question 3: Teacher beliefs related to practices of assessment. In addition to questions regarding assessment data gathering, organizing, summarizing, and reporting, the elementary general music teacher respondents were also asked four questions regarding their opinions about assessment practices. These questions specifically asked about educational rationale for assessment, unique motivators for assessment, challenges and advantages to assessment, and satisfaction level of the current assessment practices in elementary general music.

When given several reasons from which to choose, respondents were asked to choose all that applied as "educational reasons for using assessments in elementary general music." Out of the total 211 respondents in this research, 194 answered this question. When asked about using assessment to "diagnose" individual and groups, 192 and 168 respondents selected these reasons, respectively. "Assigning grades" was the

Table 25. *Beliefs Regarding Educational Reasons for Assessments.* (N = 194)

	Respondents	
Change Options	#	%
Diagnose strengths & weaknesses of students	192	99.0
Evaluate and adjust your own instruction	187	96.4
Diagnose the needs of the class as a group	168	86.6
Communicate academic progress	154	79.4
Motivate students	111	57.2
Assign grades	73	37.6
Control students	9	4.6
Other	7	3.6

option selected by 73 of the respondents. "Instructional" evaluations and adjustments were selected by 187 respondents, and 154 respondents chose "communicate academic progress." Only nine respondents selected "control students," and 111 respondents indicated using assessment to "motivate students" in the elementary general music classroom. These summaries are found in Table 25.

The seven teachers who chose "other" as an educational reason for using assessments in elementary general music also shared a short statement. The other educational reasons shared were:

- formative assessments with marker boards
- help monitor progress
- validation of subject and communication to parents of progress
- describe curricular objectives that were reached to parents
- help students become better musicians
- determine best practices; accountability and credibility
- long term instructional planning for each grade level.

These seven "other" comments were reiterations or further explanations of the given list of choices rather than totally different, unrelated rationale. All seven of the respondents who added an "other" comment had also checked other boxes or options on this question.

The second question regarding teacher opinions asked respondents about any "requirements" or "unique motivators" for classroom assessment. In other words, why do elementary general music teachers assess at all? What are the stimuli or reasons why they assess? Respondents were given numerous choices along with an "other" option.

Table 26. *Teacher Responses of Requirements or Motivators for Assessment.* (N = 200)

Change Options	Respondents #	Respondents %
Monitor student progress	173	86.5
Establish if students understand concepts	171	85.5
Motivate students	113	56.5
Allow teacher to adapt instruction	110	55.0
Provide teacher accountability	108	54.0
Assist in assigning student grades	103	51.5
Identify gifted students	96	48.0
Determine students' readiness	74	37.0
Provide validity in justifying music program(s)	74	37.0
Personal reasons	22	11.0
Certain number of assessments required by district	18	9.0
Other	4	2.0

Several options were related to student achievement; other options were about teacher adaptation and accountability. The remaining options were assigning "student grades," providing "validity in justifying music programs," a required number of assessments required by the district," and "personal reasons." The responses to this question are summarized in Table 26. The four "other" comments were: "communicate to parents," "to know if students get it or not," "none," and "achievement." Three of the four respondents who added an "other" comment had also checked other boxes or options for this question.

The next question on the survey instrument asked respondents to identify the greatest "challenges" and greatest "advantages" to assessment in elementary general

Table 27. *Respondents' Reported Greatest Challenges with Assessment.* (N = 210)

Perceived Challenges	# of Respondents
Lack of time in class	80
Lack of preparation time	62
Total number of students (large class size)	59
Total amount of data and record keeping	26
Classroom management	23
Lack of fun/joy assessment activities	14
Lack of resources	12
Variety of assessments needed	10
Differences in student population	9
Lack of training	9
Lack of technology	5
Lack of validity of assessments	4
Poor parent perception or care	4
Subjectivity or teacher opinion	4
Developing assessment tools	3
Consistency with all standards	1

Table 28. *Respondents' Reported Greatest Advantages to Assessment.* (N = 210)

Perceived Advantages	# of Respondents
Track student achievement, progress, understanding, growth	43
Improve, adjust teaching	37
Justify music class, program	8
Encourage, motivate students	7
Identify talented and gifted	7
Communicate to parents	3
Multiple reasons	3
Unsure	2
Identify standards/benchmarks	2
Grades	2
Curriculum	1

music. This was an open-ended question, with respondents encouraged to share any thoughts or comments. Tables 27 and 28 summarize a content analysis of the many responses. Several respondents shared more than one challenge or advantage. The response count, below in Tables 27 and 28, reveal the total number of respondents who indicated each response.

Finally, respondents were asked about the individual satisfaction level with the "current assessment practices, expectations, and procedures" in their teaching scenario or school setting. The choices for respondents were "strongly dissatisfied," dissatisfied," "unsure," "satisfied," and "strongly satisfied." The results of this question are summarized in Table 29.

Table 29. *Respondents' Level of Satisfaction with Assessment.* (N = 209)

	Respondents	
Satisfaction Level	#	%
Strongly Dissatisfied	11	5.3
Dissatisfied	54	25.8
Unsure	59	28.2
Satisfied	78	37.3
Strongly Satisfied	7	3.3

Of the 209 respondents, 85 teachers indicated "satisfied" or "strongly satisfied" satisfaction levels with assessment in elementary general music. Sixty-five of the respondents indicated "dissatisfied" or "strongly dissatisfied." Fifty-nine respondents indicated being "unsure" of their level of satisfaction with their current assessment

practices, expectations, and procedures in elementary general music at their current teaching scenario or school setting.

Five final questions were asked relating to respondents' opinions regarding assessment in the elementary general music classroom. A majority of the elementary music educator respondents indicated being "unsure" or "disagreeing" or "strongly disagreeing" that assessment is enjoyable. Only 68 of the 207 respondents indicated enjoying the assessment process. Furthermore, only 27 total respondents indicating enjoying the measurement issues related to assessment. A majority, 137 respondents, indicated finding "challenges" with organizing assessment data.

Table 30. *Participants' Opinions Regarding Assessment.* (N = 207)

Assessment Opinion Statement	Reported Opinions							
	SD	D	U	A	SA	M	Mdn	SD
"I enjoy assessing elem. music students"	6	67	66	57	11	3.00	3	0.96
"I enjoy the measurements issues related to assessment"	39	74	67	20	7	2.43	2	1.01
"I find many challenges when completing assessments on my students"	4	52	14	96	41	3.57	4	1.13
"There is enough preparation time to assess my elementary music students"	18	91	41	49	8	2.70	2	1.04
"The assessments I complete for elementary general music are thorough descriptions of student achievement & growth"	16	68	73	44	6	2.79	3	0.96

NOTE: Statistical analyses were abbreviated as M = mean. Mdn = Median. SD = Standard Deviation. From the survey instrument, reported opinions were abbreviated in this table as Strongly Disagree (SD), Disagree (D), Unsure (U), Agree (A) and Strongly Agree (SA).

It was only the opinion of 57 of the respondents that the quantity of preparation time outside of the classroom was adequate for assessment responsibilities and practices. Lastly, only 50 of the respondents indicated – with "agree" or "strongly agree" – that their assessment practices were "complete" and "thorough descriptions of student achievement and growth." These opinions are summarized in Table 30.

Demographics and opinions. Again, the researcher was interested in potential demographic groups of respondents that may be responding similarly to certain questions. Crosstabs (Chi-Square test of goodness of fit) were calculated between demographics and the self-reported opinions. The null hypothesis would be that the variables of demographic groups, specifically school size, number of students, number of buildings, and number of years teaching, and the opinions as shared to the afore mentioned questions regarding enjoying assessment, finding challenge with assessment, having enough preparation time, and completing thorough assessments, are independent.

As revealed in Table 31, the demographic of school size (small and large) was compared to the opinion variables of enjoying assessment, finding challenge with assessment, having enough preparation time for assessment, and completing thorough assessments in elementary general music. A Chi-Square statistic was calculated and found statistically significant between school size and respondents finding challenge with assessments practices in elementary general music, $\chi^2(1, N = 187) = 6.119, p < 0.05$. The null hypothesis would state that the variables are not dependent. This null hypothesis would be rejected when considering school size and finding challenge with assessment. An alternative hypothesis would be accepted, stating that the variables are

dependent. The other opinion variables did not have statistically significant findings with the demographic variable of school size with the Chi-Square calculations.

No statistically significant calculations were found between total number of students taught and enjoying assessment or finding challenge in assessment. As shared in Table 31, the number of students taught did not reveal significance to the opinions regarding the amount of enjoyment nor the amount of challenge found in assessment.

Table 31. *Crosstabs of Demographic Variables and Opinions.*

Variables	*Chi^2 test*
School Size	
Enjoyment of assessment	
Challenge with assessment	
Enough preparation time	
Complete thorough assessments	
Number of Students	
Enjoyment of assessment	
Challenge with assessment	
Enough preparation time	
Complete thorough assessments	
Number of Buildings	
Enjoyment of assessment	
Challenge with assessment	
Enough preparation time	
Complete thorough assessments	
Years of Teaching Experience	
Enjoyment of assessment	
Challenge with assessment	
Enough preparation time	
Complete thorough assessments	

NOTE: * $p < 0.05$ ** $p < 0.01$

Thus, the null hypothesis of no dependent variables between the total number of students variable with the enjoyment, challenge, preparation time, and thoroughness of assessments cannot be rejected.

No statistically significant relationships or associations were calculated between number of buildings served by the respondents and the respondents' opinions or answers to the opinion questions regarding assessment. With the variable of number of buildings served, then, the null hypothesis cannot be rejected. Although no significant findings were calculated, the statistics are, again, summarized in Table 31.

One statistically significant finding was calculated when the demographic variable of number of years teaching was analyzed with the reported opinions regarding the ability to complete thorough assessments in music. When comparing the respondents' number of years teaching to the respondents' reported opinions on their abilities to complete "thorough" assessments in elementary general music, a Chi-Square test calculated significant findings with $\chi^2(3, N = 198) = 24.040, p < 0.01$. Thus, a null hypothesis of no dependence between the variables could be rejected. The alternative hypothesis of the variables being dependent would be accepted for the number of years teaching and the respondents' reported abilities to complete thorough assessments in elementary general music exists. The remaining opinion variables (enjoyment, challenge and preparation time) were not found to be statistically significant.

Although not a large number of statistically significant calculations were found when comparing these demographic variables to the opinions reported by the respondents, a potential reason for lack of significance could be the overwhelming

majority of the respondents reporting "no" for many of the answers. These numbers for each demographic variable and each reported opinion are summarized in Table 32. Whereas the number of statistically significant findings were low in overall quantity, every single demographic group had a majority that reported not enjoying assessment practices, finding challenge in assessment practices, not having enough preparation time to assess, and not completing thorough assessments in elementary general music.

Table 32. *Variables Cross-Tabbed with Reported Opinions.*

Variables	Opinions							
	Enjoyment		Challenge		Prep Time		Thorough	
	NO	YES	NO	YES	NO	YES	NO	YES
School Size								
Small (A/1A/2A)	47	37	41	43	56	28	57	27
Large (3A/4A)	70	33	32	71	76	27	85	18
Number of Students								
200 or less	13	10	10	13	13	10	13	10
201 to 300	23	11	17	17	25	9	25	9
301 to 400	32	21	17	36	43	10	43	10
401 to 500	32	20	16	36	35	17	39	13
501 or above	24	9	17	16	24	9	29	4
Number of Buildings								
One	69	39	43	65	80	28	83	25
More than one	54	33	34	53	61	26	64	23
Years of Teaching								
0 to 9	59	24	25	58	61	22	72	11
10 to 19	37	14	25	26	37	14	45	6
20 to 29	24	21	20	25	32	13	26	19
30 or more	10	9	12	7	17	2	10	9

NOTE: "yes" responses are the combined "strongly agree" and "agree" responses, and "no" counts are the "unsure," "disagree," and "strongly disagree" responses.

Therefore, even without significant statistical findings, these findings are overwhelming shared by the majority of respondents throughout every demographic.

Again, in this instance, the respondents had indicated (for each of the reported opinions) a response between 1 and 5, with 3 being unsure (See Table 30). For the data analysis found in Table 32, responses of 1, 2, and 3 (strongly disagree, disagree, and unsure) were summarized as "NO" data, and responses of 4 and 5 (agree, and strongly agree) were summarized as "YES" data.

Qualitative Study: Focus Groups

The second part of this research study was qualitative inquiry, utilized to obtain further, rich data from a smaller group of participants. Focused on the four research areas of gathering, organizing, summarizing, and reporting data (Stiggins, 2005), the focus groups were positive discussions of what is really occurring in the elementary music classrooms with regard to assessment practices. The two interview groups were conveniently sampled from nearby school districts of varying sizes. One group was an urban, large district (three participants) and the second group was music teachers from nearby small districts (two participants). All participants shared freely, answered general questions and added further meaning and opinion regarding assessment topics. The participants were labeled by alphabet letters (A, B, and A, B, C) to delineate each comment, and to clarify who said what. Transcripts of the conversations were immediately transcribed and numbered by line. Participants were asked to review for accuracy. Both transcripts can be found as Appendices M and N.

With the goal of further insight into the survey questions, coding of responses into categories was the next step. Following the initial coding, the common themes that emerged from responses were: gathering assessment data (G), organizing assessment data (O), summarizing assessment data, (S), reporting assessment data (R), impacting behavioral factors (B), and beliefs regarding assessment (L). The coding abbreviations and frequencies of each in both focus groups are summarized in Table 33.

Table 33. *Focus Group Coding Categories and Frequencies.*

		Frequency	
Code	Key Concept Represented	Small School	Large School
G	Gathering Assessment Data Strategies	11	15
O	Organizing Assessment Data Strategies	3	1
S	Summarizing Assessment Data Techniques	1	5
R	Reporting Assessment Procedures	6	4
B	Behavior Factors Impacting Assessment	5	6
L	Participant Beliefs Regarding Assessment	7	15

Small School District Summary

The two volunteers in this focus group were each from small nearby school districts. One volunteer was the only elementary general music teacher in her entire district. The other volunteer had 2 elementary general music colleagues in the small district. They freely shared practices and opinions regarding assessment in each

respective small district. A summary of the coding of small school district responses is found in Appendix M.

Gathering assessment data. The coding from the Small School District music teachers described how assessment data was gathered mostly through observation. Also mentioned were singing and rhythm games where the teacher's assessment was unknown to the students. Tests were described as tools for assessing older students. Both Teacher A and Teacher B categorized assessments into two types: "observation" (lines 1 and 8), and "tests" (line 43). Rubrics were also discussed, especially used with certain units and certain grade levels (lines 65 to 70). Another method of gathering assessment data described by the teachers in small school districts was a thumbs up or thumbs down response from the younger students.

Organizing assessment data. The responses from the small school district teachers regarding organizing assessment data occurred only three times. The teachers described a portfolio where student progress is noted over time by contributing student work to a folder over the years of elementary school. The teachers agreed that portfolios were a way to see progress or lack of progress for each student. The teachers described (lines 283 and 284) having a large number of students. Teacher A taught 450 elementary general music students and Teacher B taught 373 elementary general music students.

Summarizing assessment data. The responses from the small school district teachers regarding summarizing assessment data occurred only once (line 15). Both teachers noted such a large number of students in each class and number of total classes that created a barrier in summarizing assessment data.

Reporting assessment data. Reporting assessment data was coded six times among small school district teachers. Both of the teachers indicated online "report cards" (lines 212 and 213). Teacher A described a pull-down menu of standards with rating scales on a form that was shared with both art and physical education. Additional comments are optional online also. Teacher B shared that comments can make up "for a lack of clarity" on the online forms.

Student behavior impact on assessment. Both teachers shared details about student behaviors, such as respect, attitude, and participation, and the relationship with these expected on-task behaviors to overall assessment. Teacher B described five daily points earned by each individual student every day for participation. Teacher B continued to describe points for attendance at concerts as "a big part of their assessment piece for their grade" (lines 82 to 83). Teacher A described a daily participation grade (line 113) and a balance of effort along with conceptual music skills. Teacher B described separate categories on report cards for behavior and music conceptual music skills.

Teacher beliefs regarding assessment. The opinions shared by both small school district teachers included Teacher A describing a collaborative effort with another small school district music teacher nearby. Together, as the only elementary general music teacher in their individual districts, they are collaborating on future potential assessment efforts. Teacher B reported student accountability as one advantage to assessment practices, for their work toward learning. Further, teacher accountability was also a benefit to assessment so that teachers can adjust teaching based on assessment

results. Both Teacher A and Teacher B shared that assessments reveal to the teacher student mastery, either confirming or refuting student mastery of each unique class. Again, adjusting teaching accordingly was described as a positive outcome. Both teachers further noted an extremely busy, duty-filled work week in order to work full time within their districts. Band lessons and school duties filled a large portion of their non-general music teaching time.

Large School District Summary

The three elementary general music teachers from the nearby large, urban school district also shared descriptions of assessment practices and opinions relating to music classroom assessment. Each teacher taught at a different elementary building within the district and did not travel to a multi-building assignment. The coded summaries of these volunteers' responses are found in Appendix N.

Gathering assessment data. The comments coded as Gathering Assessment Data included a description of the variation in assessment practices throughout the numerous district elementary school buildings. No uniformity in gathering data was described, with only optional opportunities to share ideas among fellow district music teachers. The large district elementary general music teachers described assessment in two ways: "observation" (lines 56 and 57) and "formal" (line 53). Teachers A and B described observation as the most commonly used method for gathering assessment data. They described "watching" students and assessing "on the fly" (line 50). Teacher A described efforts to look for students who "don't fit in" (line 57) or do not "demonstrate the standard" (line 57). Teacher B also noted listening to students sing to

assess pitch matching (line 58). Teacher A reported the use of recording devices to document each student's singing.

All three elementary general music teachers agreed that it is difficult to gather assessment data and reported difficulty in finding time to assess formally, inability to complete assessments within a 25-minute class period, difficulty in delivering assessments when students come to music directly from physical education class (lines 58 to 73). Teacher C described a difficulty in finding assessments to gather data that are also motivating to the students, specifically describing a "good lesson" (line 79) that would "allow you to assess at the same time" (line 80).

Teacher A continued to describe a large number of total elementary general music students, specifically 570, divided into 46 total sections per week. Teacher B supported that statement with a description of six elementary general music classes in a row each morning without breaks. Another barrier to gathering assessment was described by Teacher B that students do not show up for evening concerts, outside of the school day, and, thus, those performances cannot be used as a means for gathering assessment data when not all students are present. Teacher C delineated the difference between knowing about music and performing music, and how these two different skills should both be assessed. These barriers to gathering assessment data were described with great detail.

Organizing assessment data. In addition to gathering assessment data, the large school district music teachers also described their efforts of organizing assessment data. Due to the lack of time and successful strategies to gather data, Teacher B

described a primarily subjective summary of student progress. With no assessment data gathered, the need to organize was nonexistent (lines 141-151).

Summarizing assessment data. The large district elementary general music teachers described an online system for summarizing assessment data in preparation to report to parents. The computer system was further described as only a comment box, with no standards listed, unlike the other elementary core subjects in the district. Assessment data summarization efforts were further described by Teacher B, when she described having typed detailed comments about each student into the computerized system, only to find out that those comments were not shared with parents. This was found as "frustrating" (line 178) whereas Teacher B chose not to make comments after that point in any future summarizing attempts.

Reporting assessment data. Reporting practices of the large district general music teachers were described as "on the computer" with "no other input" from special area teachers (lines 25 and 26). Teacher B described how parents of the elementary students do not "come to the music room during conferences" (line 191) and how typically there was no contact with most parents other than the computerized comments.

Student behavior impact on assessment. The large district elementary general music teachers stated complications in separating behaviors from conceptual skills when assessing students. Teacher B described how misbehaviors created barriers to assessment whereas when students are not participating or if they are off task, then the assessments are not "true" (line 104). Teacher B has students who do not sing or participate. These students are constantly off task and Teacher B spends more time "managing the

behaviors" than assessing. Teacher B questioned the legitimacy of the classroom assessment efforts due to the repeated bad behaviors of the students. Teacher B questioned if any grade reported was truly based on students' "actual skill level" (line 108) or "based on a behavior that gets in the way of demonstrating the skill" (lines 108 and 109).

Teacher C added similar comments of students "not doing a good group effort" (line 110) and are not participating or "doing the skill" (line 111). Teacher C described a 3-2-1 rating scale for a "participation grade" (line 113) for each class but not any assessment of musical skills. Teacher C continued to share the difficulty in separating behaviors from true conceptual assessment.

Teacher A described how behavior, specifically "effort and participation" (line 173), were formerly separated on a previous reporting system. The skills and behaviors were to be graded individually. The current computerized reporting system, however, does not have conceptual skills separate from behaviors in the music classroom.

Teacher B continued that effort and participation highly influence her music reporting, as "trying" (line 194) was described as "half the skill right there" (line 196). Teacher A disagreed with this approach, and stated that behaviors should not influence grading. Teacher A indicated that "participation" (line 209) could be included in the grade but "behavior" (line 209) could not. Teacher C reiterated the difficulty in assessing a musical concept if a student is not participating in music class. This lack of participation by the student does not provide "work" (line 213) for the teacher to assess.

Teacher belief regarding assessment. All three large district music teachers shared several beliefs related to assessment and reporting in elementary general music. Teachers A and B agreed that one advantage to teaching in a large district was having consultation opportunities other professionals with regard to assessment. Teacher C described assessment "advice" (line 16) she received from colleagues in the district as beneficial.

Both Teachers A and B shared their beliefs of being overwhelmed by the large number of total music students, extra building duties required, and the rigorous schedule of teaching so many sections of each grade in the large district. Teacher A stated that, due to the large number of students, it is very difficult to even recognize names and faces of the first year students in kindergarten. Teacher C further described inequity among the multiple district elementary buildings in in expectations, procedures, and teaching assignments that was problematic in the large district.

Teachers A and C both noted the other accommodations required when teaching in this large district. They noted students with learning accommodations, such as students who need things read to them, students needing one-to-one attention, students needing workspace away from distractions, and students needing someone to write for them. These specific accommodations were described as limitations when attempting to use any type of written assessment in the elementary classroom. Furthermore, other physical accommodations were also shared, such as hearing impairments, wheelchairs, and students with physical limitations. These were described as "road block[s]" (line 127) and inhibitors to assessment.

When specifically describing the overwhelming beliefs associated with assessment, Teacher C stated, “there are just too many things we have to think about sometimes” (line 81). Teacher B used the words “I was frustrated” (line 182) when her reporting efforts were futile.

Summaries of Quantitative and Qualitative Data

As revealed in both the quantitative data and the qualitative data, the respondents in both the survey (N = 211) and the focus groups (N = 5) had large amounts of information to share regarding assessment in elementary general music. The researcher utilized data to describe the demographics of the respondents and to answer the three research questions within this study.

Demographics

To describe the demographic variables of the quantitative section of this research project, descriptive statistics including measures of central tendencies were utilized. The survey respondents were from all sizes of schools in Iowa, from A to 4A, with the most respondents in the 4A category. All of the Area Educational Agency regions throughout the state were also represented by the survey respondents. The respondents also indicated how many years they had taught elementary general music. Two respondents were in their first year of teaching. The most veteran respondent had taught elementary music for 38 years. The average amount of teaching experience was 14.1 years. A large majority (121) of the respondents held bachelor’s degrees, and 77 respondents reported having master’s degrees. One respondent indicated having a doctorate. Approximately half of the survey respondents (109) served one building, and the other respondents indicated

traveling to 2, 3, or 4 buildings. Over 90% of the respondents taught Kindergarten through fourth grade general music, with many others also teaching fifth and sixth grades. Only 26.3% of the respondents reported teaching pre-school music. The respondents were evenly divided in total number of students taught, with a majority of respondents (61.7%) having class sizes of 21 to 25 students. The participants in the focus groups were conveniently selected from districts nearby the researcher with two individuals from small districts, both with less than five years of teaching experience, and three individuals from a large district, one veteran teacher and two others with less than seven years of teaching experience.

Research Question 1: Gathering, Organizing, Summarizing, Reporting

To answer research question number one, descriptive statistics were also used. This question investigated the types and frequencies of assessments practices that were reported by the respondents in elementary general music classrooms in Iowa. The first question was designed to gather data regarding gathering, organizing, summarizing, and reporting data in elementary music. Assessment gathering practices were reported by survey respondents as primarily observation (96.7%), recordings (30.8%), and concert performances (29.9%). Focus group members also indicated observation as the primary means of gather assessment data. Almost all of the survey respondents indicated assessing singing (99.5%) and rhythm (99.0%) objectives. All but nine of the survey respondents indicated a lack of class time to assess student skills and knowledge. The Focus group participants also reported observation as the primary assessment gathering

tool, and also noted many inhibiting circumstances to gathering data in the music classroom.

Organizational assessment practices were summarized by survey respondents as being “limited” (166 respondents). Focus group participants supported this data, when they shared that the large number of total students taught were limiting to their organizational capabilities. Only 84 (42.6%) of the survey respondents indicated transferring data into any type of organized system such as a spreadsheet or any computerized organizational tool. When organizing data, survey respondents did report that multiple curricular areas were considered, such as singing, rhythm, melody, form, harmony, instruments, listening, movement, and music appreciation. Both of the two small-school focus group participants described a method of organizing the assessment data they gathered, using a seating chart and writing down, using a rating scale, what is observed for each student. One of the small-school participants also described a folder where each student’s work is kept over time. This portfolio-type scenario was described as a tool for creating a means to see the progress of the students.

When considering the summarizing of data, the researcher sought efforts or attempts by elementary general music teachers to take data and draw conclusions or summaries for indications of achievement or mastery by students, and also to make decisions for future teaching, for remediation or moving forward in the curriculum. Approximately half, or 108 (54.8%) survey respondents, reported having analyzed class data or data taken across grade levels and seeking trends for moving forward or reinforcing instruction. The focus groups supported these claims when they stated having

looked, through observation, for students that looked confused and in need of further instruction. Both survey respondents and focus group participants indicated utilizing other attributes beyond curricular goals when summarizing assessment data. Criteria such as participation, attitude, effort, and behavior were reported as being used as part of assessment summarization.

When reporting assessment data, 171 of respondents indicated using report cards. Only 28 of survey respondents indicated the use of letter grades, however, 168 respondents indicated a similar categorical rating scale. Approximately two-thirds of the survey respondents stated that their school district mandated a report card format, with the other respondents indicating using self-authored materials. Most shared a common form throughout the year, with slight variations for each grade level. Both the small-school and large-school focus group participants indicated using an online format for report cards, with very little opportunity for flexibility, comments, or variability. Survey respondents and focus group participants both indicated the required expectation of being present during parent conferences. A majority (60%) of the survey respondents reported being visited by parents to discuss student progress during conference times. The focus group participants, however, indicated a highly unlikely occurrence of parents visiting during conferences. Of the survey respondents, only 42 (N = 201) indicated that they collected enough data to accurately communicate student growth. Furthermore, only 50 respondents (N = 201) reported the ability to "provide adequate documentation" if or when a parent would question a student's assessment in music.

Research Question 2: Demographic Implications

The second research question examined if the demographic data of the respondents would have any significance with the reported assessment practices of gathering, organizing, summarizing, and reporting data in elementary general music. Statistically significant relationships were discovered using the Chi Square calculation. A significant finding was discovered between the total number of students taught and practices in gathering data. Respondents indicated inabilities to successfully gather data in the elementary general classroom when teaching large number of students. The number of years of teaching was a second demographic variable found to have statistical significance with assessment practices. Significant findings occurred between years of teaching and gathering, organizing, and summarizing assessment data.

The focus group data further supported the potential relationship between the total number of students and the ability to gather assessment data. Both the small school group and the large school group indicated inabilities to gather assessment data due to large class sizes and a large total number of students. The small-school focus group participants described an ideal data gathering activity where the students were playing a game all while the teacher was assessing their skills through participation in the game. The teacher observed each student's abilities as they participate and makes note of the quality or achievement level of the students.

Research Question 3: Related Opinions

Finally, to answer research question number three, respondents were asked in both the survey instrument and in focus group discussions about beliefs or opinions

related to assessment in elementary general music. Both the respondents' survey data and the focus group members' comments (of small and large schools) revealed how lack of class time and lack of preparation time outside the classroom were both high concerns with regard to assessment practices. The quantitative data further revealed the respondents' main reasons for using assessment practices in elementary general music were diagnostic in nature, for student achievement and also to adjust teaching as necessary.

Approximately 60% of the respondents surveyed indicated a lack of satisfaction with assessment practices overall. The focus groups triangulated this finding, as well, with the respondents having indicated low satisfaction. Furthermore, enjoyment was also reported very low in the survey respondents and the focus groups. Many challenges were identified by the survey respondents and the focus group participants as lack of preparation time, lack of training, a busy schedule and large numbers of students.

With regard to potential relationships between demographic groups within the respondents and their opinions, a Chi Square statistic was calculated and statistically significant differences were found between the size of respondents' school and increased challenge with assessment, and also as the total number of years teaching increased, so did the respondents' reported abilities to complete thorough assessments.

Conclusion

The data from this mixed method research was vast and thoroughly described what was occurring in elementary general music assessment in Iowa. The next chapter,

Chapter 5, presents the interpretations, implications, and suggestions for improvement in the future.

CHAPTER 5

CONCLUSIONS

The purpose of this study was to research the practices in assessment of elementary general music teachers in Iowa. Using the four assessment categories from Stiggins (2005), research questions were based on gathering evidence, organizing or storing evidence, summarizing evidence, and sharing or reporting evidence of student assessment in elementary general music. A thorough literature review was completed with a multitude of information relating to assessment in general and assessment in the elementary general music classroom. A survey (Appendix F) was designed and used to gather data from elementary music teachers throughout the state. A total of 211 elementary general music teachers responded to the survey request, either online or in writing. Two focus groups were also formed to gather further, rich background information in support of the three research questions. The focus groups were volunteers from surrounding small districts (2 members) and a nearby large district (3 members). The summaries, discussion, implications and recommendations in this chapter are based on the conclusions from the literature review, the survey data, and the focus group data.

Limitations

This study was conducted of the current population of Iowa's public and private school elementary general music teachers. The data gathered in survey responses did not necessarily represent overall responses of all elementary music teachers in the United States. Findings should not be generalized beyond logical parameters. The validity of

this study was dependent on the accuracy of responses provided by the respondents that choose to volunteer.

Delimitations

This study focused on the assessment and reporting practices of elementary general music teachers in Iowa. This study did not include middle school or high school music, nor did it include any elementary instrumental ensembles in the schools. No considerations for differences in gender or ethnicity among teachers were considered. No considerations for any unique makeup of student populations in music classrooms were considered.

The survey was distributed to a representative sample of Iowa. In addition, all of the music teachers in the three counties surrounding the researcher also received the survey. The two focus groups were selected from these three surrounding counties for convenience. All data was limited only to the degree to which teachers expressed their practices and beliefs.

Data Summaries and Discussion

This study was designed to answer three research questions relating to the assessment practices of elementary general music teachers in Iowa. This research study was designed to collect data specifically addressing questions.

1. What are the features of the classroom assessment practices currently implemented to:
    a. gather evidence (assessment data) of student learning in elementary general music in Iowa?

b. organize student learning data in elementary general music in Iowa?

c. summarize assessment data in elementary general music in Iowa?

d. report assessment data in elementary general music in Iowa?

2. What demographics, or teacher characteristics, influence or relate to classrooms assessment practices of gathering, organizing, summarizing, and reporting assessment data in elementary general music in Iowa?
3. What are teacher beliefs or opinions related to assessment in elementary general music in Iowa?

Each of these questions will be addressed with data from the survey responses and focus group informational data.

Demographics of Study Respondents and Participants

To describe the demographic variables of the quantitative section of this research project, descriptive statistics including measures of central tendencies were utilized. The survey respondents were from all sizes of schools in Iowa, from A to 4A, with the most respondents in the 4A category. All of the Area Educational Agency regions throughout the state were also represented by the survey respondents. The respondents also indicated how many years they had taught elementary general music. Two respondents were in their first year of teaching. The most veteran respondent had taught elementary music for 38 years. The average amount of teaching experience was 14.1 years. A large majority (121) of the respondents held bachelor's degrees, and 77 respondents reported having master's degrees. One respondent indicated having a doctorate. Approximately half of the survey respondents (109) served one building, and the other respondents indicated

traveling to 2, 3, or 4 buildings. Over 90% of the respondents taught Kindergarten through fourth grade general music, with many others also teaching fifth and sixth grades. Only 26.3% of the respondents reported teaching pre-school music. The respondents were evenly divided in total number of students taught, with a majority of respondents (61.7%) having class sizes of 21 to 25 students.

The participants in the focus groups were conveniently selected from districts nearby the researcher. Two elementary general music teachers from small districts agreed to participate. Both of the participants in the small-school focus group possessed less than five years of teaching experience. One of the participants was from a rural setting, where two towns had merged into a very small district. This participant was the only music teacher for both of the small elementary schools in the district. The other small-school focus group participant was from a very small school within a larger community. Three individuals from a large district volunteered to participate. In the large-school focus group, one participant was a veteran teacher and the two others had less than seven years of teaching experience each. All three participants were from the same large, urban district, with all three of them assigned to only one building within the large district.

Research Question 1

The first research question was written to gather evidence to describe the specific types and frequency of assessment practices being used in Iowa elementary general music classrooms. Using Stiggins' (2005) summation of assessment practices, Question 1 was designed to inquire about the gathering, organizing, summarizing, and reporting practices

in elementary general music. Descriptive statistics were utilized to summarize data to address how respondents gathered, organized, summarized, and reported assessment data in elementary music.

Assessment gathering practices. Assessment gathering practices were reported by survey respondents as primarily observation (96.7%), recordings (30.8%), and concert performances (29.9%). The number of assessments increased as the elementary general music students got older, indicating that the respondents are not assessing the younger students as often as the older students. Almost all of the survey respondents assessed singing (99.5%) and rhythm (99.0%) objectives. All but nine of the survey respondents indicated a lack of class time to assess students' musical skills and knowledge. One of the themes that emerged from my analysis of the focus group participants was that the "majority" (large-school line 56) of their data gathering was through observations. They specifically stated how they were "looking" (large-school line 57) for student behaviors. Furthermore, one large-school participant described data gathering practices as "watching students" (large-school, line 54). Both of the small-school participants described how data gathering practices were commonly games where students had to demonstrate by singing or performing a rhythm. The survey respondents and focus group participants both indicated a wide use of observation as their primary means of gathering assessment data.

The researcher was astounded at the high frequencies or percentages of respondents who utilized informal assessment, such as observations, in the elementary general music room. Albeit, assessment is challenging – especially in an expressive and

performance-based subject like music. But much of music teaching at the elementary level can be quantified and, thus, able to be documented and assessed formally in writing. Curricular items such as notes on a staff, intervals, note durations, and many terminology definitions could be formally assessed in writing in a worksheet or testing format. Furthermore, assessment tools such as these exist and can be found (free of charge) online. The low reported occurrence of these formal types of assessment was unanticipated. Rationale for these self-reported actions was documented as lacking in class time with students. Further explanation was also large class sizes and total number of students. Many respondents also indicated a very grueling teaching schedule with very little breaks in the day. An overall lack of preparation time was also reported. With these barriers, it is understandable why more formal assessing is not occurring in the elementary general music classrooms across Iowa.

The use of concert attendance as an assessment tool was also unexpected by the researcher. With elementary students, the sole responsibility of showing up and participating in an evening performance would lie with parents or guardians who freed up the child's schedule and provided the transportation to the concert. Thus, penalizing students for concert absence seemed to be a questionable practice. Granted, concerts may provide an opportunity for the music teacher to observe and evaluate all students performing at once, but it could be impossible for parents or guardians to physically get students to school – at no fault of the student. This researcher disagreed with using concerts with any aspect of assessment in elementary general music whereas attendance would not have a direct connection to skills or knowledge mastery. A self-reflection after

a performance by the students could be a potential opportunity for brainstorming means for improvement. Again, students who did not attend the concert would not be able to complete this assessment, but not because of their own lack of concept mastery.

Assessment organizational practices. Organizational assessment practices were summarized by survey respondents as being "limited" (166 respondents). Only 84 (42.6%) of the survey respondents indicated transferring data into any type of organized system such as a spreadsheet or any computerized organizational tool. When organizing data, survey respondents did report that multiple curricular areas were considered, such as singing, rhythm, melody, form, harmony, instruments, listening, movement, and music appreciation. Both of the two small-school focus group participants described a method of organizing the assessment data they gathered, using a seating chart and writing down, using a rating scale, what is observed for each student. One of the small-school participants also described a folder where each student's work is kept over time. This portfolio-type scenario was described as a tool for creating a means to see the progress of the students.

An interesting factor when considering organizational practices of the respondents is the lack of sharing of assessment tools used in the assessment process. One aspect of this research study was the optional opportunity for respondents to submit documents or forms used with assessment practices. Only one respondent, out of 211 total respondents, emailed the researcher with any type of form, design, program, or rubric used with assessment practices. The unknown factor is if these items existed and were not shared, or if they did not exist.

Assessment summarizing practices. When considering the summarizing of data, the researcher sought efforts or attempts by elementary general music teachers to take data and draw conclusions or summaries for indications of achievement or mastery by students, and also to make decisions for future teaching, for remediation or moving forward in the curriculum. Approximately half, or 108 (54.8%) survey respondents, reported having analyzed class data or data taken across grade levels and seeking trends for moving forward or reinforcing instruction.

Survey respondents also indicated utilizing other attributes beyond curricular content when summarizing assessment data. Criteria such as participation, attitude, effort, and behavior were reported as being used as part of assessment summarization process. Clarification was not made, however, how these non-curricular criteria contributed toward a finalized summation of musical achievement or performance. The researcher asked, in two different questions, the amount of impact that non-curricular criteria should have in assessment. All but 59 of the respondents indicated a positive outcome for students that "tried hard but performed poorly" when it came to music objectives. Thus, the other 142 respondents on this question (N = 201) agreed that effort was just as important – if not more important – than content mastery. When asked the second time, only 36 respondents indicated that they would only consider curricular content – with the other 162 respondents (N = 198) having indicated that when students "tried hard but performed poorly" on music objectives, it was satisfactory.

Due to this reported situation, follow up questions were asked in both focus groups and discussed later in this chapter. The small-school focus group participants

both agreed that behavioral, or non-curricular criteria, have a place within assessment practices. One participant described a "balance" (small-school line 240) of effort and actual skill. The other small-school participant noted behavioral options for reporting on the district online reporting system. One large-school participant stated the difficulty in separating behavior from skill, and another large-school participant questioned the authenticity of skills assessment as students are so far off task and not participating that it was impossible to assess.

The summarizing of assessment data in elementary general music rooms is most definitely variable. With no standard forms or documentation to use, teachers have found difficulty in summarizing or making conclusions about such varied – or lack of any – data. The issue of what criteria – curricular only or including non-curricular – to use within assessments is troubling. Many teachers are factoring in students' behaviors, especially off-task and inappropriate behaviors, into assessments. Other teachers are summarizing behaviors, but in a separate part of assessment summarization. Other teachers, still, do not figure student behaviors into summarizations at all. This calls to question the validity of elementary general music assessments as a whole in the state of Iowa. The same students could receive varying summaries of elementary music – depending on the teacher, building, or district.

Assessment reporting practices. When reporting assessment data, 171 of respondents indicated using report cards. Only 28 of survey respondents indicated the use of letter grades, however, 168 respondents indicated a similar categorical rating scale. Approximately two-thirds of the survey respondents stated that their school district

mandated a specific, district-wide report card format, while the other respondents indicating using self-authored materials. Most shared a common form throughout the year, with slight variations for each grade level. This reveals that students throughout Iowa may receive, depending on the district or school, a hand-written note from a music teacher or an official form with a letter grade (A through F). This also reveals that students and parents may receive identical copies of the same form, numerous times throughout the year, regardless of what content has been taught or covered during the different times of the year. Furthermore, students and parents may receive communication that is identical, regardless of the grade-level of the student. This would leave room for very little specification within each content area. For example, rhythm would be a general concept, rather than specifically indicating rhythms of mastery (example: quarter notes, dotted half notes, etc.).

Survey respondents and focus group participants both indicated the required expectation of being present during parent conferences. Both the small-school and large-school focus group participants indicated using an online format for report cards, with very little opportunity for flexibility, comments, or variability.

This data indicated, again, a lack of consistency across respondents, buildings, and districts in Iowa. The volunteers from the focus groups described an online computer-generated reporting system whereas survey respondents indicated paper forms, some of which were original and others district mandated. Many respondents indicated that parents visited with them during school conferences, where others did not. The most alarming data collected in this research project was when over 75% of the survey

respondents indicated an inability to rationalize assessment reporting practices if ever questioned. This illustrated a continued trend of lack of continuity in data gathering, organizing, summarizing and reporting assessment data altogether.

Furthermore, and even more alarming, was the even smaller number of respondents who claimed the inability to gather or use data to report or communicate student progress to parents or administration. This data was not particular to any particular demographic group, but rather spread across all demographics (See Table 16 in Chapter 2). Out of 201 respondents, 159 of them indicated not collecting "enough data to accurately communicate student growth." In other words, approximately 79% of the responding elementary general music teachers in Iowa cannot confidently assess student growth in the music classroom. Furthermore, 149 of the same respondents reported not being able to "provide adequate documentation" for parents with questions regarding assessment. Again, this is a startlingly high percentage of elementary general music teachers in Iowa who do not feel competent or confident in assessment practices.

The data also revealed potential reasons for these assessment practices not to be occurring. Large class sizes, overall total number of students, demanding schedules, and student behaviors were the primary reasons shared by both the survey respondents and the focus group participants as inhibitors to assessment practices.

Research Question 2

The second research question was designed to investigate potential relationships between certain respondent demographic groups and the answers to research Question 1 and 3. Not only was this research project designed to describe the assessment practices,

but determine if any groups of respondents that may be responding in similar or dissimilar ways.

Respondents indicated great inabilities to gather data in the elementary general classroom when teaching larger number of students. This was consistent with no regard to any other demographic (school size, number of buildings served, years of teaching experience). This could be considered a rational conclusion as larger numbers of students would require larger amounts of assessment-related work for elementary music teachers. It would sensibly be more difficult to gather, organize, summarize and report data from larger class sizes and larger student body populations than from smaller groups of students. It is not surprising to the researcher that large numbers of students would create much difficulty in working with assessment data in elementary general music classes.

The focus group data further supported the potential relationship between the total number of students and the ability to gather assessment data. Both the small school group and the large school group indicated inabilities to gather assessment data due to large class sizes and a large total number of students. The large school group specifically noted the difficulty of assessing so many students with such little time. Furthermore, with multiple classes in each teaching day and no breaks in between classes, it is impossible to remember the assessment data of so many students by the end of the day. The large school participants also noted a sizeable number (46) of sections in each week and the difficulties associated with gathering data for so many students. These, too, are logical summations when dealing with hundreds of elementary music students and assessments. As many of the surveys and focus group discussions described gathering

methods as observation, this would require teachers to have individual students share musical experiences and the notation of the experiences. Another method of data gathering reported was a written instrument. The design, printing, distribution, completion, retrieving, and correcting of these tools would also require much more time with higher numbers of students. Understandably, smaller total numbers of students would make the management of data much simpler.

The number of years of teaching was a demographic variable found to have statistical significance with some assessment practices. Significant findings occurred between years of teaching and organizing and summarizing assessment data. A Chi-Square test was calculated and found to be significant between years of teaching experience and organizing assessment data, $\chi^2(4, N = 188) = 12.862$, $p < 0.01$, and between years of teaching experience and summarizing assessment data, $\chi^2(4, N = 188) = 22.665$, $p < 0.01$.

These significant findings were not surprising to the researcher. Assessment practices are, like many skills, refined over time. It is logical that as respondents' teaching experience increased, the more likely they were to find success in gathering, summarizing and analyzing assessment data from students while teaching elementary general music. The teaching experience could be attributed to gaining expertise through years of practice, having repeated assessment practices over time and becoming more proficient and adept through repetitions over time. Another logical explanation may be that with time and experience come abilities to deal with adversity and to overcome difficult challenges. Even as assessment practices and expectations change through the

years, experienced teachers become more and more skilled at dealing with change and with stressful situations in the classroom.

The focus group volunteers were of varying years of teaching experience. All of the members, regardless of years of teaching experience, shared difficulties in gathering, organizing and summarizing assessment data. In fact, the volunteer with the most years of teaching expressed "difficulties" (Large school, line 135) in differentiating assessment gathering tasks. Thus, the focus group discussions do not reinforce the idea of a relationship between years of teaching experience and gathering assessment data, but rather that all teachers – with no connection to years of teaching – experience challenge in working with assessment data.

The focus group also agreed on finding challenge specifically with organizing and summarizing assessment data, again with no relationship to years of teaching experience. In fact, one teacher surmised no need for organizing or summarizing data, as no data existed – since none was collected! The point that was made was that so many barriers exist when organizing and summarizing assessment data in elementary music, that the tasks seem unsurmountable. These thoughts were not unique to inexperienced teachers, but rather shared among all of the focus group participants regardless of years of teaching experience. The researcher concluded that any organizing or summarizing that was occurring was haphazard and unregulated.

Research Question 3

The third research question was designed to discover any opinions, attitudes, or ideas that could influence or contribute to the practices to assessment in Iowa elementary

general music. By asking for opinions about and attitudes toward assessment, the researcher was seeking rationale or explanation for practices, or lack of practices, in elementary general music.

Quantitative and qualitative data. All but two of the respondents agreed that diagnosing student strengths and weaknesses was a viable reason for assessment in elementary general music. Only seven respondents did not agree that assessment could also be used for adjusting teaching to better instruct the students. These responses are related and sensible to the researcher as it would be desirable for teachers to use assessment data to reflect on how teaching could be changed or improved to better deliver material for increased retention or achievement. Nine of the respondents indicated that assessments were used to "control students." This was a shocking finding by the researcher that teachers would use assessments as some type of punishment, threat, or modifier of behaviors in the classroom.

When asked about levels of satisfaction with assessment practices, survey respondents shared their opinions with 85 or 46% (N=209) indicating satisfaction. With only a minority indicating satisfaction, further questions were asked about perceived advantages and challenges in elementary music assessment. Eighty total respondents indicated two main advantages: diagnosing student achievement or progress, and improving or adjusting teaching. These reported advantages correspond to the responses in earlier questions about rationale or reasons for assessment. The perceived challenges in elementary music were primarily reported as lacking in class time, lacking in

preparation time, and large numbers of students. These are logical responses, considering the number of tasks necessary to assess in music.

Overall, a majority of the survey respondents indicated not enjoying the tasks associated with assessment in elementary music nor did they enjoy the measurement issues and terms related to assessment. A large majority of the respondents indicated finding many challenges when completing assessments in music, and an even larger majority (all but 50) sharing that the assessments that are completed are not a thorough description of student achievement and growth.

These concerns, again, are understandable to the researcher as so many of the respondents have very large numbers of students. In fact, as the number of students taught increased, so did the respondents responses that revealed lacking time to prepare assessments outside of class. A further reasonable finding was the correlation between total number of students and their reported inability to complete thorough assessments on students. Again, this finding is sensible considering the teaching scenario and challenges associated with sheer high numbers of students.

The focus group conversations reiterated the survey data with similar opinions and additional explanations and detail. With 570 students in 46 sections of teaching elementary general music each week and with the added element of little or no breaks in each teaching day, it is no surprise to find results that indicate finding challenge in all assessment practices. The large district group of teachers further explained how within the large groups of students are numerous students with behavioral needs and academic

accommodations. The steps necessary for these accommodations prevented one focus group member from attempting assessment at all.

When statistical calculations were used, the researcher was seeking tendencies of demographic groups in relationship to the respondents' reported opinions. A statistically significant correlation was calculated between the school size and finding challenge with assessment practices. A Chi-Square statistic was calculated and found to be $\chi^2(1, N = 187) = 6.119$, $p < 0.05$. These findings were also reinforced by the focus group data, as the large school volunteers expressed much more concern with overall concern and challenges with assessment. Many comments from these participants related to misbehaviors in the classroom and the inability to manage the classroom while assessing students. The small school participants did not explain similar challenges.

The most significant findings were calculated between the opinions reported by respondents and the number of years of teaching of each respondent. When the number of years teaching were compared to whether or not the respondents felt that they completed thorough assessments with elementary general music students, a Chi-Square statistic was found to be $\chi^2(3, N = 198) = 24.040$, $p < 0.01$. This indicates that the responses of the teachers who took the survey did not answer the question about completing thorough assessments as would be expected throughout the sample.

These findings may be considered reasonable when the researcher, again, contemplated the many valuable traits that emerge in teachers with each year of teaching experience. The understandings gained with teaching experience would, understandably, add to the level of enjoyment through, if nothing else, mere familiarity and practice over

time. Decreasing challenges would also be a logical occurrence with experience and practice. Do younger, less experienced music teachers have the teaching abilities to master grueling schedules and large class sizes, and difficult student behaviors all while also assessing the students' achievement, growth, and skills?

Even though few statistically significant findings were discovered with relationship to the demographic groups of respondents, that, in itself, is a significant finding. As revealed in Table 30 of Chapter 2, the reason why significance was not found with relation to respondents' opinions was that they all – regardless of most all demographics – were responding similarly: that majorities were not finding enjoyment, were finding challenge, did not report enough preparation time, and did not feel they were thoroughly completing assessments. The significance here is that these self-reported opinions were not specific to any one demographic, but rather common among a large majority of all respondents. These opinions are startling – across all demographics. The majority of elementary music teachers do not enjoy assessment, find challenge with assessment, do not feel they have enough preparation time, and do not feel that they complete thorough assessments of their students. These issues are endemic to all of the music teachers who participated in this study.

Relationship with Previous Research

These findings are consistent with previous research, as reported in Chapter 2 of this study. Again, with reference to the original research questions, these findings are broadly in line with previous research studies.

Research Question 1: Gathering, Organizing, Summarizing, and Reporting

My findings on gathering assessment data relate similarly to research studies of Hartwell (1979), Miller (1990), Hepworth-Osiowy (2004), where observation prevailed as the dominant means of gathering assessment data. Similar to Rasor's 1998 study, the respondents in my study focused on singing as a major attribute in assessment. Contrary to the literature, my study also revealed a high likelihood of elementary music teachers to assess beat and rhythm. The previous research studies did not focus in on specific content areas, but rather revealed large variation in content that was assessed. When summarizing data, this study found very similar findings to previous studies (Carter, 1986; Rasor, 1988; Hepworth-Osiowy, 2004; Barkley, 2006) that many non-academic factors were considered in assessment summaries, in addition to the academic content and concepts.

Also prevalent in the literature (Nightingale-Abell, 1993; Talley, 2005; Barkley, 2006), and in line with my findings were the high occurrences of informal assessments such as observation and a lack of formal assessments (Talley, 2005), such as exams and commercial tests. Furthermore, many former research studies similarly concluded a vast variation among teachers when identifying means of assessment (Carter, 1986; Niebur, 1997; Shih, 1997; Hepworth-Osiowy, 2004).

Research Question 2: Demographic Influences on Assessment

Again, the findings in this study were not unlike those revealed in the literature review of previous studies. Anderson-Nickel (1997) also found a difference in assessment practices between beginning and more experienced elementary music

teachers. Also in line with the literature (Nightingale-Abell, 1993; Niebur, 1997; Hepworth-Osiowy, 2004; Salvador, 2011) this study found majorities of elementary music teachers find challenge with assessment, regardless of demographic criteria.

Research Question 3: Teacher Opinions and Beliefs

Similar to the literature findings (Hartwell, 1979; Nightingale-Abell, 1993; Barkley, 2006; Salvador, 2011), this study concurred that teachers have strong feelings about assessment, especially the inhibitors that prohibit them from excelling at assessment. Common to the literature and to this study were opinions regarding lack of release or preparation time, lack of training, hectic teaching schedules, and large numbers of students in music classes.

Implications and Recommendations

Based on the literature review and this research study, many implications and recommendations have been discovered. Both the literature and this study showed that the most common assessment practices by elementary general music teachers in Iowa were observations, recordings, and attendance or participation at concerts. The least common assessment practices were standardized music ability and aptitude tests. Overall, elementary general music teachers in Iowa indicated observation and performance-based assessment. More experienced teachers found more ease in many assessment practices. A majority of the respondents and participants found great challenge in gathering, organizing, summarizing, and reporting assessment data. Large majorities did not feel that assessment practices were completed thoroughly nor could they substantiate their findings to parents. The literature and this study also indicated that

music teachers were teaching large sections of students with very limited preparation time.

The study appears to support the argument for change and several potential implications or suggestions for elementary general music teachers and districts for future improvement.

1. Increased enjoyment and completion of assessment practices in elementary general music was found with more experienced music teachers. Therefore, school districts should find means of in-servicing or instructing lesser experienced music teachers in assessment practices. Furthermore, Iowa could implement a mentoring system in music education specifically for assessment purposes.
2. Similarly, the majority of respondents in this study, as well as reported in the literature, expressed challenge in assessment practices. In order to find more enjoyment in assessment, elementary music teachers could take additional courses or professional development in assessment.
3. There was a multitude of reports of large class sizes in music. If assessment is an expected part of elementary general music education, school districts could provide opportunities for smaller class sizes.
4. There was a plethora of reports of large total numbers of students taught. If detailed and thorough assessments are expected as part of elementary general music education, school districts may provide a maximum total number of students taught by any one elementary music teacher.

5. There was a large majority of elementary general music teachers who reported utilizing observation as a means of gathering assessment data. Districts, states, or publishers may want to create or provide standard forms for organizing observational data specific to elementary general music education. Technological apps might also provide an efficient and convenient manner for elementary music teachers, perhaps with names, grades, and outcomes prepared for recording performances or electronically converting observations to summarization and ready for reporting.
6. The literature also suggested integrating assessment tasks into activities (Niebur, 1994). This would be ideal for a smooth continuation of teaching content while also gathering assessment data. One of the small-school focus group participants also described this as an over-arching goal for assessment.
7. Although this researcher supports local curricular decision making practices, this study revealed large varieties in assessment practices. Due to this, perhaps districts or states may consider uniform or standard expectations and/or reporting forms or formats in elementary general music assessment practices.

Suggestions for Future Research

The results of this research study provided a summary of elementary general music teachers' assessment practices in Iowa. The study, however, had several things that could be improved upon in possible replication.

1. The researcher included questions on the survey that were not utilized in analysis. These questions, thus, could be eliminated in future research. Questions on future surveys should relate directly to research questions.
2. There were no significant findings among elementary general music teachers who taught in one building and those that traveled to two or more buildings. Further questions regarding traveling details could be asked about accommodations for traveling teachers that were not discovered in this study. These accommodations could have an impact on assessment practices.
3. Many respondents did not answer several questions as revealed in the varying N's for each question. Any steps to encourage participants to answer every question could improve possibilities for analysis.
4. When designing survey questions and survey responses, future researchers must consider what types of data is necessary for certain types of statistical analysis.
5. With the large numbers of respondents having indicated lack of enjoyment and lack of thorough completion of assessments, future researchers could investigate the amount of higher education institutions that offer assessment coursework at the undergraduate level in music education.

Conclusion

This study provided much insight into the assessment practices of elementary general music teachers in Iowa. With the large numbers of observations that occurred in elementary general music classrooms, perhaps future research or publications will

include potential means of notating large numbers of observational data in an efficient and timely manner as well as the organizing and summarizing of observation notes. Until educational priorities shift away from large class sizes and demanding music teaching schedules, elementary music teachers will continue to find many challenges with assessment practices. Future studies in the area of elementary general music research could give elementary general music teachers even more practical information on how to improve assessment practices in elementary general music. This research would be beneficial in all states, not just Iowa, to see if these challenges and variations are common nationwide. If this is, indeed, the case, then serious consideration would need to occur as to whether or not assessment in elementary general music is valued, valid, reliable, and worthwhile.

REFERENCES

Abeles, H. F. (2010). Assessing music learning. In H. F. Abeles & L. A. Custodera (Eds.), *Critical issues in music education: Contemporary theory and practice* (pp. 167-193). New York, NY: Oxford University Press.

Abrahams, F. E. (2000). Implementing the national standards for music education in pre-service teacher education programs: A qualitative study of two schools. (Doctoral Dissertation: Temple University). *Dissertations Abstracts International, 61*(03), 0925A.

Adams, T. L., & Hsu, J. W. Y. (1998). Classroom assessment: Teachers' conceptions and practices in mathematics. *School Science and Mathematics, 98,* 174-180. doi:10.1111/j.1949-8594.1998.tb17413.x

Airasian, P. W. (1991). Perspectives on measurement instruction. *Educational Measurement: Issues and Practice, 10*(1), 13-16. doi:10.1111/j.1745-3992.1991.tb00172.x

Airasian, P. W. (1994). *Classroom assessment* (2nd ed.). New York, NY: McGraw-Hill.

Airasian, P. W. (2000). *Assessment in the classroom: A concise approach*. Boston, MA: McGraw-Hill Higher Education.

Airasian, P. W., Kellaghan, T., Madaus, G. F., & Pedulla, J. (1977). Proportion and direction of teacher rating changes of pupil progress attributable to standardized test information. *Journal of Educational Psychology, 69*(6), 702-7-9. doi:10.1037/0022-0663.69.6.702

Airasian, P. W., & Russell, M. (2008). *Classroom assessment: Concepts and applications*. Boston, MA: McGraw-Hill Higher Education.

Allen, J. D. (1983, April). *Classroom management: Students' perspectives, goals and strategies.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning, *Journal of Educational Strategies, Issues, and Ideas, 78*(5), 218-224. doi:10.3200/TCHS.7.5.218-223

Allen, J. D., & Lambating, J. (2001, April). *Validity and reliability in assessment and grading: Perspectives of preservice and inservice teachers and teacher education professors*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Ambach, G. M. (1996). Music education, arts standards, and student performance. In MENC (Ed.), *Aiming for excellence: The impact of the standards movement on music education* (pp. 5-13). Reston, VA: MENC.

Anderson, L. W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making.* Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson-Nickel, J. D. (1997). Teacher expertise among elementary general music teachers. (Doctoral Dissertation: Arizona State University). *Dissertations Abstracts International, 58*(6), 2127.

Asmus, E. P. (1999). Music assessment concepts. *Music Educators Journal, 86*(2), 19-24.

Austin, S., & McCann, R. A. (1992). *Here's another arbitrary grade for your collection: A statewide study of grading policies.* Philadelphia, PA: Research for Better Schools.

Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how. In T. R. Guskey (Ed.), *ASCD yearbook 1996: Communicating student learning* (pp. 19-40). Alexandria, VA: Association for Supervision and Curriculum Development.

Barkley, M. (2006). Assessment of the national standards for music education: A study of elementary general music teacher attitudes and practices. (Doctoral Dissertation: Wayne State University). *ProQuest Dissertations & Theses A&I*, 1439697.

Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education, 36*(4), 46-49. doi:10.1177/ 002248718503600412

Bateson, D. J. (1990). Measurement and evaluation practices of British Columbia science teachers. *Alberta Journal of Educational Research, 36,* 46-49.

Bell, C. L. (2003). Beginning the dialogue: Teachers respond to the national standards in music. *Bulletin of the Council for Research in Music Education, 156,* 31-42.

Bentley, A. (1966). *Measures of musical abilities.* London, England: George A. Harrap.

Berman, S. (2008). *Performance-based learning: Aligning experiential tasks and assessment to increase learning.* Thousand Oaks, CA: Corwin Press.

Black, P. J. (1998). *Testing, friend or foe? The theory and practice of assessment and testing.* London, England: Falmer Press.

Boeije, H. (2009). *Analysis in qualitative research.* Thousand Oaks, CA: Sage Publications.

Boothroyd, R. A., & McMorris, R. F. (1992). *What do teachers know about testing and how did they find out?* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

Bouton, K. (2001). What does a grade of S, N, or U mean to parents? In MENC (Ed.), *Spotlight on assessment on music education*. Reston, VA: MENC – The National Association for Music Education.

Bower, G. H., & Hilgard, E. R. (1981). *Theories of Learning*. Englewood Cliffs, N.J.: Prentice-Hall.

Boyle, J. D. (1996). The national standards: Some implications for assessment. In MENC (Ed.), *Aiming for excellence: The impact of the standards movement on music education* (pp.109-115). Reston, VA: MENC.

Boyle, J. D. (compiler for National Commission on Instruction). (1974). *Instructional objectives in music: Resources for planning, instruction and evaluating achievement*. Vienna, VA: Music Educators National Conference.

Boyle, J. D., & Radocy, R. E. (1987). *Measurement and evaluation of musical experiences*. New York, NY: Schirmer Books.

Bracey, G. W. (2000). *Bail me out! Handling difficult data and tough questions about public schools.* Thousand Oaks, CA: Corwin Press.

Branscome, E. (2005). A historical analysis of textbook development in American music: Education and the impetus for the national standards for music education. *Arts Education Policy Review, 107,* 13-19. doi:10.3200/AEPR.107.2.13-20

Britzman, D. P. (1991). *Practice makes practices*. New York, NY: State University of New York Press.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice, 10,* 35-36.

Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*(2), 123-142. doi:10.1111/j.1745-3984.1993.tbo1080.x

Brookhart, S. M. (1994). Teachers' grading: Practices and theory. *Applied Measurement in Education, 7*(4), 279-301. doi:10.1207/S15324818ame0704-2

Brookhart, S. M. (2004a). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record, 106*(3), 429-458. doi:10.1111/j.1467.9620.2004.00346.x

Brookhart, S. M. (2004b). *Grading*. Upper Saddle River, NJ: Pearson Education.

Brophy, T. S. (2000). *Assessing the developing musician: A guide for general music teachers.* Chicago, IL: GIA Publications.

Brophy, T. S. (2003). Developing and implementing standards-based assessments. In C. Lindeman (Ed.), *Benchmarks in action: A guide to standards-based assessment in music* (pp. 1-16). Reston, VA: The National Association for Music Education.

Brophy, T. S. (Ed.) (2008). *Assessment in music education: Integrating curriculum, theory, and practice.* Chicago, IL: GIA Publications.

Brophy, T. S. (Ed.) (2010). *The practice of assessment in music education: Frameworks, models, and designs. Proceedings of the 2009 Florida Symposium on Assessment in Music Education*, University of Florida. Chicago, IL: GIA Publications.

Brummett, V. M. (1993). The development, application, and critique of an interactive student evaluation framework for elementary general music. (Doctoral Dissertation: Universtiy of Illinois at Urbana-Champaign). *Dissertations Abstracts International, 58*(03), 1590.

Brummett, V. M., & Haywood, J. (1997). Authentic assessment in school music: Implementing a framework. *General Music Today, 11*(1), 4-10.

Burbridge, A. A. (2001). Assessment: Pencil, paper... & performance, too! In MENC (Ed.), *Spotlight on assessment in music education*. Reston, VA: MENC – The National Association for Music Education.

Butler, S. M., & McMunn, N. D. (2006). *A teacher's guide to classroom assessment: Understanding and using assessment to improve student learning*. San Francisco, CA: John Wiley & Sons.

Byo, S. J. (1997). General education classroom teachers' and music specialists' perceived ability to implement the national standards for music education. *Dissertations Abstract International, 85*(02), 0409A.

Byo, S. J. (1999). Classroom teachers' and music specialists' perceived ability to implement the national standards for music education. *Journal of Research in Music Education, 47*(2), 111-123. doi:10.2307/3345717

Calfee, R. C., & Masuda, W. V. (1997). Classroom assessment as inquiry. In G. D. Phye, (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 69-102). San Diego, CA: Academic Press.

Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *Journal of Educational Research, 93*(6), 350-356. doi:10.1080/00220670009598729

Carlin, J. (1996). Videotape as an assessment tool. *Teaching Music, 3*(4), 38-39.

Carr, J. F., & Harris, D. E. (2001). *Succeeding with standards: Linking curriculum, assessment, and action planning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Carter, K. (1984). Do teachers understand principles for writing tests? *Journal of Teacher Education, 35*, 57-60.

Carter, K. G. (1986). The status of vocal/general music programs in Oklahoma elementary schools. (Doctoral Dissertation: University of Oklahoma). *Dissertations Abstracts International, 47,* 3349A.

Chase, C. I. (1999). *Contemporary assessment for educators*. New York, NY: Addison-Wesley Educational Publishers.

Chiodo, P. (2001). Assessing a cast of thousands, *Music Educators Journal, 87*(6), 17–23. doi:10.2307/3399687

Chiodo, P., Frakes, L., MacLeod, S., Pagel, R., Shuler, S., Thompson, J., & Watson, D. (1998). Making the grade: Authentic assessment in music, K-8. *Teaching Music, 6*(2), 34035; 48.

Chittenden, E. (1991). Authentic assessment, evaluation, and documentation of student performance. In V. Perrone (Ed.), *Expanding student assessment. Association for Supervision and Curriculum Development* (pp. 22-31). Alexandria, VA: Association for Supervision and Curriculum Development.

Cizek, G. J. (1996). Grades: The final frontier in assessment reform. *NASSP Bulletin*, 103-110. doi:10.1177/019263659608058416

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 1-32). San Diego, CA: Academic Press.

Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3*, 159-179. doi:10.1207/S15326977ea0302_3

Colwell, R. J. (1969). *Music achievement tests 1 and 2*. Chicago, IL: Follett Educational.

Colwell, R. J. (1970). *Music achievement tests 3 and 4*. Chicago, IL: Follett Educational.

Colwell, R. J. (1974). Musical achievement: Difficulties and directions in evaluation. In J. D. Boyle (Ed.), *Instructional objectives in music: Resources for planning instruction and evaluating achievement* (pp. 127-135). Vienna, VA: Music Educators National Conference.

Colwell, R. J. (1979). *Silver Burdett Music Competency Tests*. Morristown, NJ: Silver Burdett Company.

Colwell, R. J. (1995). Now that we're there, who wants to know? *Quarterly Journal of Music Teaching and Learning, 6*(4), 5-18.

Colwell, R. (2002). Assessment's potential in music education. In R. Colwell & C. Richardson (Eds), *The new handbook of research on music teaching and learning: A project of the Music Educators National Conference*. New York, NY: Oxford University Press.

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

Colwell, R. (2010). Many voice, one goal: Practices of large-scale music assessment. In T. S. Brophy (Ed.), *The practice of assessment in music education: Frameworks, models and designs* (pp. 3-22). Chicago: GIA Publications.

Consortium of National Arts Education Associations. (1994). *National standards arts education: What every young American should know and be able to do in the arts*. Reston, VA: MENC - The National Association for Music Education.

Cope, C. (1996). Steps toward effective assessment. *Music Educators Journal 83*(1), 39-42.

Cope, D. (1999). One approach to musical intelligence. *IEEE Intelligent Systems, 14*(3), 21-25.

Council of Chief State School Officers [CCSSO]. (1997). *Guidelines for arts assessment.* Arts Education Consensus Project. Retrieved from http://www.state.nj.us/education/njpep/classroom/arts_assessment/worddocs/GuidelinesforArtsAssessments.ppt.

Cross, L. H., & Frary, R. B. (1996, April). *Hodgepodge grading: Endorsed by students and teachers alike.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, New York, NY.

Cross L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education, 12*(1), 53-72. doi:10.1207/S15324818ame1201_4

Davidson, B. J. (1995). Leadership styles of successful male and female college choral directors. (Doctoral Dissertation: Arizona State University). *Dissertations Abstracts International, 56*(12), 4692. (UMI No. 9611661).

Denzin N. K., & Lincoln Y. S. (2003). *Collecting and interpreting qualitative materials* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2nd ed.). New York, NY: John Wiley & Sons.

Dorr-Bremme, D. W. (1983). Assessing students: Teachers' routine practices and reasoning. *Evaluation Comment, 6*(4), 1-12.

Dorr-Bremme, D., & Herman, J. (1986). *Assessing student achievement: A profile of classroom practices* (CSE Monograph Series in Evaluation No. 11). Los Angeles, CA: University of California, Center for the Study of Evaluation.

Drake, R. M. (1957). *Drake musical aptitude tests*. Sarasota, FL: Raleigh M. Drake.

Dubois, P. H. (1970). *A history of psychological testing.* Boston, MA: Allyn & Bacon.

Duerksen, G. L. (1995). Using planned assessment techniques to increase student growth in music education. *Quarterly Journal of Music Teaching and Learning, 6*(4), 57-64.

Duncan, S. A. (2009). *Assessment practices of string teachers*. Master of Music Thesis, University of Miami. Retrieved from http://etd.library.miami.edu/theses/available/etd-03252009-172116/unrestricted/duncans2Sp09.pdf.

Education Commission of the States. (1983). *A nation at risk.* Retrieved from http://www.ed.gov/pubs/NatAtRisk/risk.html.

Edwards, N., & Richey, H. G. (1947). *The school in the American social order*. Cambridge, MA: Houghton Mifflin.

Elford, G. W. (2002). *Beyond standardized testing: Better information for school accountability and management.* Lanham, MD: Scarecrow Press.

Ellis, J. F. (1963). Should children receive achievement grades in music? *Canadian Music Educator, 4*(2), 29.

Farnum, S. E. (1969). *Farnum music tests*. Riverside, RI: Bond Publishing.

Farrell, S. R. (1997). *Tools for powerful student evaluation: A practical source of authentic assessment strategies for music teachers*. Galesville, MD: Meredith Music Publications.

Fautley, M. (2010). *Assessment in music education*. New York, NY: Oxford University Press.

Fiese, R. K., & Fiese, R. E. (2001). Music teaching and assessment. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 13-15). Reston, VA: MENC – The National Association for Music Education.

Foley, B. (2001). Performance assessment: Application for general music. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 16-17). Reston, VA: MENC – The National Association for Music Education.

Fransen, J. C. (1998). An investigation into the use of portfolio assessment in elementary music education. (Masters Thesis, University of Manitoba). *Dissertation Abstracts International*.

Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice, 12*(3), 23-30. doi:10.1111/j.1745-3993.1993.tb00539.x

Friedman, S. J., & Frisbie, D. A. (2000). Making report cards measure up. *The Education Digest, 65*(5), 45-50.

Gallagher, J. D. (1998). *Classroom assessment for teachers*. Upper Saddle River, NJ: Prentice-Hall.

Gallavan, N. P. (2009a). *Developing performance-based assessments: Grades K-5*. Thousand Oaks, CA: Corwin Press.

Gallavan, N. P. (2009b). *Developing performance-based assessments: Grades 6-12.* Thousand Oaks, CA: Corwin Press.

Ginsburg, M. B. & Clift, R. T. (1990). The hidden curriculum of preservice teacher education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 450-465). New York, NY: Macmillan.

Giordano, G. (2007). *How testing came to dominate American schools: The history of educational assessment.* New York, NY: Peter Lang Publishing.

Glickman, C. (1993). *Renewing America's schools*. San Francisco, CA: Jossey-Bass.

Goolsby, T. (1995). Portfolio assessment for better evaluation. *Music Educators Journal, 82*(3), 39-44.

Goolsby, T. (1999). Assessment in instrumental music. *Music Educators Journal 86*(2), 31-35, 50. doi:10.2307/3399587

Gordon, E. E. (1965). *Musical aptitude profile*. Boston, MA: Houghton Mifflin.

Gordon, E. E. (1970). *Iowa tests of music literacy*. Iowa City, IA: Bureau of Educational Research and Service, University of Iowa.

Gordon, E. E. (1979). *Primary measures of music audiation.* Chicago, IL: G.I.A. Publications.

Gordon, E. E. (1982). *Intermediate measures of music audiation.* Chicago, IL: BG.I.A. Publications.

Gordon, E. E. (1991). *Iowa tests of music literacy: Revised*. Chicago, IL: G.I.A. Publishers.

Gordon, E. E. (2000). *Jump right in: The music curriculum.* Chicago, IL: G.I.A. Publishers.

Gordon, E. E. (2007). *Learning sequences in music: A contemporary music learning theory*. Chicago, IL: GIA Publications.

Gredler, M. W. (1999). *Classroom assessment and learning*. New York, NY: Longman.

Green, L. (2001). Evaluating creativity: Making and learning by young people. *British Journal of Music Education, 17*(3), 303-316.

Greer, R. D. (1980). *Design for music learning.* New York, NY: Teachers College Press.

Griswold, P. A., & Griswold, M. M. (1992, April). *The grading contingency: Graders' beliefs and expectations and the assessment ingredients.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Gronlund N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York, NY: Macmillan.

Gullickson, A. R. (1982, November). *The practice of testing in elementary and secondary school*. Paper presented at the Rural Education Conference, Manhattan, KS. (ERIC Document Reproduction Service No. ED 229391).

Gullickson, A. R. (1985). Student evaluation techniques and their relationship to grade and curriculum. *Journal of Educational Research, 79*(2), 96-100.

Guskey, T. R. (1994). Making the grade: What benefits students? *Educational Leadership 52*(2), 14-20.

Guskey, T. R. (1996). Reporting on student learning: Lessons from the past - prescriptions for the future. In T. R. Guskey (Ed.), *ASCD yearbook 1996: Communicating student learning* (pp. 13-24). Alexandria, VA: Association for Supervision and Curriculum Development.

Guskey, T. R. (2001). Helping standards make the GRADE. *Educational Leadership, 59*(1), 20-27.

Guskey, T. R. (2002). *How's my kid doing? A parent's guide to grades, marks, and report cards*. San Francisco, CA: John Wiley & Sons.

Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press.

Haladyna, T. M. (1999). *A complete guide to student grading.* Boston, MA: Allyn & Bacon.

Haladyna, T., Haas, N., & Allison, J. (1998). Continuing tensions in standardized testing. *Childhood Education, 74*(5), 262-273.

Hale, C. L., & Green, S. K. (2009). Six key principles for music assessment. *Music Educators Journal, 95*(4), 27-31.

Hamann, K. L. (2001). Assessment tools for the music classroom. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 23-25). Reston, VA: MENC – The National Association for Music Education.

Hanzlik, T. J. (2001). An examination of Iowa high school instrumental band directors' assessment practices and attitudes toward assessment (Doctoral Dissertation: The University of Nebraska-Lincoln). *Dissertation Abstracts International, 62,* 955A.

Harrison, L. N. (1983). *Getting started in* elementary music education. Englewood Cliffs, NJ: Prentice-Hall.

Hartwell, R. W. (1979). The status of elementary music education in the public schools of the state of Ohio: Implications for educational change. (Doctoral Dissertation: University of Cincinnati). *Dissertations Abstract International, 40*(07), AAT 8002111.

Hepworth-Osiowy, K. (2004). Assessment in elementary music education: Perspectives and practices of teachers in Winnipeg Public Schools. (Master's Thesis, University of Manitoba). *Dissertation Abstracts International*.

Herman, J. L., & Dorr-Bremme, D. W. (1982, March). *Assessing students: Teachers' routing practices and reasoning.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Herrold, R. M. (1991). *New approaches to elementary classroom music* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Hickey, M. (1999). Assessment rubrics for music composition. *Music Educators Journal, 85*(4), 36-33, 52.

Hill, C. (2008). A portfolio model for music educators. *Music Educators Journal, 95*(10), 61-72.

Hill, K. W. (1999). A descriptive study of assessment procedures, assessment attitudes, and grading policies in selected public high school band performance classrooms in Mississippi (Doctoral Dissertation, The University of Southern Mississippi). *Dissertation Abstracts International, 60,* 1954A.

Hoffer, C. R. (1993). *Introduction to music education*. Belmont, CA: Wadsworth.

Hoffer, C. R. (2008). Issues in the assessment of K-12 music instruction. In T. Brophy (Ed.), *Assessment in music education: Integrating curriculum, theory, and practice* (pp. 29-36). Chicago, IL: GIA Publishers.

Hoffer, M. L., & Hoffer, C. R. (1987). *Music in the elementary classroom: Musicianship and teaching*. New York, NY: Harcourt Brace Jovanovich.

Holt-Reynolds, D. (1992). Personal history-based beliefs as relevant prior knowledge in coursework: Can we practice what we preach? *American Educational Research Journal, 29*(2), 325-349. doi:10.3102/00028312029002325

Hoolsema, M. (2010). Assessment in music education. *Music Educators Journal, 97*(1), 18-19.

Howe, F., & Simmons, B. J. (2005). *Nurturing the parent-teacher alliance*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Iowa Department of Education (2012). *Iowa SINA schools*. Retrieved from https://www.educateiowa.gov/sites/files/ed/documents/2012-2013%20Iowa%20SINA%20Schools%20%28Oct.%202012%29.pdf

Iowa High School Athletic Association. (2013). *School classification and enrollment information.* [Data file]. Retrieved from http://www.iahsaa.org/wp-content/uploads/2014/08/2013-2014_Enrollments_Numerical_revised-8-29-14.pdf

Iowa High School Music Association. (2012). *Resources.* [Data file]. Retrieved from http://ihsma.org/resources/

Jindrich, J. (1996). Empowering students through assessment. In B. Harp (Ed.), *The handbook of literacy assessment and evaluation* (pp. 11-18). Norwood, MA: Christopher-Gordon Publishers.

Johnson, R. (2002). *Using data to close the achievement gap.* Thousand Oaks, CA: Corwin Press.

Johnson, T. W., & Reed, R. F. (Eds.) (2002). *Historical documents in American education*. Boston, MA: Allyn & Bacon.

Johnson, J., Dupuis, V., Musial, D. Hall, G., & Gollnick, D. (2002). *Introduction to the foundations of American education*, (12th. Ed.) Boston, MA: Allyn & Bacon.

Kancianic, P. M. (2006). Classroom assessment in United States high school band programs: Methods, purposes, and influences. (Doctoral Dissertation: University of Maryland, College Park). *ProQuest Dissertations Publishing.* 3222315.

Keenan-Takagi, K. (2000). Embedding assessment in choral teaching. *Music Educators Journal, 86*(4), 42-46, 63. doi:10.2307/3399605

Kellaghan, T., Madaus, G. F., & Airasian, P. W. (1982). *The effects of standardized testing*. Boston, MA: Kluwer-Nijhoff.

Kelly, S. N. (2001). Using portfolios for assessing performing ensemble. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 26-28). Reston, VA: MENC – The National Association for Music Education.

Kelsey, K. D. (2001). Overcoming standardized testing with authentic assessment strategies in the classroom. *The Agricultural Education Magazine, 73*(5), 41.

Killian, J. (1995, October). *Assessment in music performing groups: Some practical suggestions*. Paper presented at the Suncoast Music Education Forum, Tampa, FL. Retrieved from http://music.arts.usf.edu/smef/art-jk.htm

Kirschenbaum, H., Napier, R., & Simon, S. B. (1971). *Wad-ja-get? The grading game in American education*. New York, NY: Hart Publishing.

Kriske J, & DeLelles, R. (2005). *Gameplan grade 1: An active music curriculum for children*. Las Vegas, NV: KiD sounds.

Kotora, E. J. (2001). Assessment practices in the choral music classroom: A survey of Ohio high school choral music teachers and college choral methods teachers. (Doctoral Dissertation: Case Western Reserve University). *Dissertations Abstracts International, 62*(12), 4102A. (UMI No. 3036343).

Kotora, 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*, 65-80.

Labuta, J. A. (1974). *Guide to accountability in music instruction*. West Nyack, NY: Parker Publishing.

Lake, K., & Kafka, K. (1996). Reporting methods in grades K-8. In T. R. Guskey (Ed.), *ASCD yearbook 1996: Communicating student learning* (pp. 90-118). Alexandria, VA: Association of Supervision and Curriculum Development.

Lavender, C. (2000). *The ultimate music assessment and evaluation kit for the general music teacher: Practical strategies, time-saving tips, samples, forms and more*. Milwaukee, WI: Hal Leonard.

Law, L. N.C. (2012). Assessing and understanding individual differences in music perception abilities. (Doctoral Dissertation, The University of York, 2012). *ProQuest Dissertations Publishing.* U603467, 1512387079.

Lehman, P. R. (1968). *Tests and measurements in music*. Englewood Cliffs, NJ: Prentice-Hall.

Lehman, P. R. (1989). Assessing your program's effectiveness. *Music Educators Journal, 74*(4), 26-29.

Lehman, P. R. (1992). Assessing learning in the music classroom. *Bulletin of the National Association of Secondary School Principals, 76*(544), 56-62.

Lehman, P. R. (1993). Implication of national standards. *Music Educators Journal, 80*(3), 25-28.

Lehman, P. R. (1996). Introduction. In MENC (Ed.), *Aiming for excellence: The impact of the standards movement on music education* (pp.1-2). Reston, VA: MENC.

Lehman, P. R. (1997). Assessment and grading. *Teaching Music, 5*(3), 58-59.

Lehman, P. R. (1998). Grading practices in music: A report of the Music Educators National Conference. *Music Educators Journal, 84*(5), 37-40.

Lehman, P. R. (2000). The power of the National Standards for Music Education. In B. Reimer (Ed.), *Performing with understanding: The challenge of the National Standards for Music Education* (pp. 3-9). Reston, VA: MENC - The National Association for Music Education.

Leonhard, C., & House, R. W. (1972). *Foundations and principles of music education*. New York, NY: McGraw-Hill.

Lincoln, E. A., & Workman, L. L. (1936). *Testing and the uses of test results.* New York, NY: Macmillan.

Lindeman, C. A. (Ed). (2003). *Benchmarks in action: A guide to standards-based assessment in music.* Reston, VA: The National Association for Music Education.

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching*, (8th ed.) Englewood Cliffs, NJ: Merrill.

Lomax, R. G. (1996). On becoming assessment literate: An initial look at preservice teachers' beliefs and practices. *Teacher Educator, 31*(4), 292-303. doi:10.1080108878739609555122

Long, N. H. (1970). *Indiana-Oregon music discrimination tests.* Bloomington, IN: Midwest Music Tests.

Lopez, C. (2001). Assessing elementary improvisation. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 32-34). Reston, VA: MENC – The National Association for Music Education.

Lortie, D. (1975). *School teacher*. Chicago, IL: University of Chicago Press.

Louk, D. P. (2002). National standards for music education: General music teachers' attitudes and practices. (Doctoral Dissertation: Arizona State University). *Dissertation Abstract International, 63* (02), 0485A.

Manke, M. P., & Loyd, B. H. (1990, April). *An investigation of non-achievement related factors influencing teachers' grading practices*. Paper presented at the National Council on Measurement in Education, Boston: MA.

Mark, M. (1995). Music education and the national standards: An historical review. *The Quarterly, 6*(2), 34-43.

Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J. (2006). *Classroom assessment and grading that work.* Alexandria, VA: ASCD.

Mason, J. (2002). *Qualitative researching* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Matheny, J. (1994). A grading system for secondary music groups. *Music Educators Journal, 81*(1), 37-39. doi:10.2307/3398729

May, G. (2001). Performance-based assessment and increased student performance. In MENC (Ed.), *Spotlight on assessment on music education* (pp. 35-36). Reston, VA: MENC – The National Association for Music Education.

McClung, A. C. (1996). A descriptive study of learning assessment and grading practices in the high school choral music performance classrooms (Doctoral Dissertation, The Florida State University). *Dissertations Abstracts International, 57*, 3437A.

McClung, A. C. (1997). Learning assessment and grading practices in the high school choral music performance classroom. *Southeastern Journal of Music Education, 9,* 252-272.

McClung, A. C. (2000). Extramusical skills in the music classroom. *Music Educators Journal, 86*(5), 37-42. doi: 10.2307/3399635

McCoy, C. W. (1988). An exploratory study of grading criteria among selected Ohio ensemble directors. *Contributions to Music Education, 15*, 15-19.

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. doi:10.2307/3344718

McCreary, T. J. (2001). Methods and perceptions of assessment in secondary instrumental music. (Doctoral Dissertation, University of Hawaii at Manoa). *ProQuest Dissertations Publishing*, 3030187.

McFarlin, K. (1965). Criteria for assigning grades. *The Instrumentalist, 22*(11), 73-74.

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*, 20-32. doi:10.1111/j.1745-3992.2001.tb00055.x

McMillan, J. H., & Lawson, S. R. (2001). *Secondary science teachers' classroom assessment and grading practices.* Metropolitan Education Research Consortium, Richmond, VA. (ERIC Document Reproduction Service N. Ed 450 158).

McMunn, N., Schenck, P., & McColskey, W. (2003, April). *Standards-based assessment, grading, and reporting in classrooms: Can district training and support change teacher practice?* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

McPherson, G. (1993). Evaluating improvisational ability of high school instrumentalists. *Bulletin of the Council for Research in Music Education, 199*, 11-20.

McTighe, J., & Ferrara, S. (1998). *Assessing learning in the classroom*. Annapolis Junction, MD: National Education Association.

Mierzwik, D. (2005). *Classroom record keeping made simple: Tips for time-strapped teachers.* Thousand Oaks, CA: Corwin Press.

Miles, M. B., & Huberman A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Miller, C. C. (2005). *Assessment in the music classroom: Practical tips and activities for measuring student progress.* Milwaukee, WI: Hal Leonard.

Miller, M. C. (1990). *The development of an affective behavior assessment instrument for use with third-grade children in a music class situation* (Doctoral Dissertation: The Florida State University). *ProQuest Dissertations & Theses A&I*, 9103105.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. Upper Saddle River, NJ: Pearson Education.

Monroe, R. (1995). Music education in Ohio high schools? A comparative study of the opinions of music teachers, principals, and professors of music education. (Doctoral Dissertation: Miami University). *Dissertations Abstracts International*, 56-01A, 16760.

Music Educators National Conference. (1974). *The school music program: Description and standards.* Reston, VA: MENC.

Music Educators National Conference. (1994a). *National standards for arts education: What every young American should know and be able to do in the art.* Reston, VA: MENC.

Music Educators National Conference. (1994b). *Perspectives on implementation: Arts education standards for America's students.* Reston, VA: MENC.

Music Educators National Conference. (1994c). *Opportunity-to-learn standards for music instruction: Grades PreK-12.* Reston, VA: MENC.

Music Educators National Conference. (1994d). *The school music program: A new vision.* Reston, VA: MENC.

Music Educators National Conference. (1996a). *Aiming for excellence: The impact of the standards movement on music education*. Reston, VA: MENC.

Music Educators National Conference. (1996b). *Performance standards for music grades prek-12: Strategies and benchmarks for assessing progress toward the national standards.* Reston, VA: National Association for Music Education.

Music Educators National Conference. (2001). *Spotlight on assessment in music education: Selected articles from state MEA journals.* Reston, VA: MENC – The National Association for Music Education.

Music Educators National Conference. (2007). *October 2007 report of the MENC task force on national standards: Report of the MENC national executive board.* Retrieved from http://www.menc.org/resources/view/october-2007-report-of-the-menc-task-force-on-national-standards.

Music Educators National Conference. (2010). *Performance standards for music: Assessment strategies for music.* Retrieved Nov. 7, 2010, from http://www.menc.org/resources/view/performance-standards-for-music-introduction.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington D.C.: United States Government Printing Office.

National Council on Education Standards and Testing. (1992). *Raising standards for American education: A report to Congress, the Secretary of Education, the National Education Goals Panel, and the American people.* Washington D. C.: United States Government Printing Office.

Nava, F. J. G., & Loyd, B. H. (1992, April). *An investigation of achievement and non-achievement criteria in elementary and secondary school grading*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Newman, D. C. (1981). Teacher competency in classroom testing, measurement preparation, and classroom testing practices. (Doctoral Dissertation: Georgia State University). *Dissertations Abstracts International, 45(*2), 1111A.

Niebur, L. (1994). Assessment as a class activity. *Music Educators Journal, 80*(5), 23-25. doi:10.2307/3398744

Niebur, L. L. (1997). Standards, assessment, and stories of practice in general music. Doctoral Dissertation: Arizona State University). *Dissertations Abstracts International, 58*(03), 791.

Nierman, G. E. (1996). Assessment - Applying the yardstick to musical growth. In MENC (Ed.), *Spotlight on assessment on music education*. (2001, pp. 40-41). Reston, VA: MENC – The National Association for Music Education.

Nierman, G. E. (1997). Tools for assessing musical skills. In MENC (Ed.), *Spotlight on assessment on music education.* (pp. 45-48). Reston, VA: MENC – The National Association for Music Education.

Nightingale-Abell, S. E. (1993). Teacher evaluation practices in the elementary general music classroom: A study of three teachers (Doctoral Dissertation: College Conservatory of Music at the University of Cincinnati). *Dissertations Abstracts International, 55*(04), 900A.

Nitko, A. J. (1996). *Educational assessment of students.* Englewood Cliffs, NJ: Prentice Hall.

Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.) Upper Saddle River, NJ: Merrill.

Nutter, K. (1999). Managing assessment in music education. *Teaching Music, 7*(1), 26-31.

O'Connor, K. (1999). *How to grade for learning*. Arlington Heights, IL: SkyLight Professional Development.

Office of Technology Assessment (1992). *Testing in American schools: Asking the right questions, OTA-SET-519.* Washington, DC: U.S. Government Printing Office, U.S. Congress.

Olson, G. H. (1989, March). *On the validity of performance grades: The relationship between teacher-assigned grades and standard measures of subject matter acquisition.* Paper presented at the Annual meeting of the National Council of Measurement in Education, San Francisco, CA.

Olson, L. (1995). Cards on the table. *Education Week, 15*(41), 23-28.

Oosterhof, A. (2001). *Classroom applications of educational measurement* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Osborne, N. (1983). The making of grades. *International Journal of Music Education, 2,* 53-58.

Orman, E. K. (2002). Comparison of the national standards for music education and elementary music specialists' use of class time. *Journal of Research in Music Education, 50*(2), 155-164.

Ornstein, A. C. (1994). Grading practices and policies: An overview and some suggestions. *NASSP Bulletin, 78*(561), 55–64.

Pace, R. D. (1972). Grading system leads to success. *The Instrumentalist, 27*(3), 20.

Page, E. B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. *Journal of Educational Psychology, 49*, 173-181. doi:10.1037/h0041940

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332. doi:10.2307/1170741

Parsons, T. (1959). The school class as a social system: Some of its functions in American society. *Harvard Educational Review, 29*, 297-318.

Perrone, V. (1977). *The abuses of standardized testing.* Bloomington, IN: Phi Delta Kappa Educational Foundation.

Phye, G. D. (Ed.) (1997). *Handbook of classroom assessment: Learning, achievement, and adjustment.* San Diego, CA: Academic Press.

Pilcher-Carlton, J., & Oosterhof, A. (1993, April). *A case study analysis of parents', teachers', and students' perception of the meaning of grades: Identification of discrepancies, their consequences, and obstacles to their resolution*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.

Pohl, M. (2000). *Learning to think, thinking to learn: Models and strategies to develop a classroom culture of thinking.* Cheltenham, Australia: Hawker Brownlow.

Pontious, M. (2001). Grading, instruction, and assessment in music. In MENC (Ed.), *Spotlight on assessment on music education* (pp. 60-62). Reston, VA: MENC – The National Association for Music Education.

Popham, W. J. (2001). *The truth about testing: An educator's call to action.* Alexandria, VA: Association for Supervision and Curriculum Development.

Power B. M., & Chandler, K. (1998). *Well-chosen words: Narrative assessments and report card comments*. York, MN: Stenhouse Publishers.

Radocy, R. E. (1986). On quantifying the uncountable in musical behavior. *Bulletin of the Council for Research in Music Education, 88*, 22-31.

Radocy, R. E. (1989). Evaluating student achievement. *Music Educators Journal, 74*(4), 30-33.

Radocy, R. E. (1995). Planning assessment in music education: It's not risk free. *The Quarterly Journal of Music Teaching and Learning, 6*(4), 19-25.

Radocy, R. E., & Boyle, J. D. (1988). *Psychological foundations of musical behavior* (2nd ed.). Springfield, IL: Charles C. Thomas Publisher.

Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education, 26*(7), 1872-1380. doi:10.1016/j.tate.2010.03.008

Rasor, S. H. (1988). A study and analysis of general music education, K-8, in the public schools of Ohio. (Doctoral Dissertation: University of Cincinnati). *Dissertations Abstract International, 50*(02). AAT 8908457

Riley, P. E. (2009). Pre-service music educators' perceptions of the national standards for music education. *Visions of Research in Music Education, 14*. Retrieved from http://www-usr.rider.edu/~vrme/v14n1/vision/Riley%20Final2.pdf. ISSN: 1938-2065.

Robinson, M. (1995). Alternative assessment techniques for teachers. *Music Educators Journal, 81*(5), 28-34. doi:10.2307/3398853

Roeber, E. D. (1995). *Using new forms of assessment to assist in achieving student equity: Experiences of the CCSSO state collaborative on assessment and student* standards. Boston, MA: Kluwer Academic Publishers.

Rosen, K. (1999). To sit beside: Informing classroom practice through the teaching, learning, assessment process. (Ph.D. Dissertation: The Claremont Graduate University). *Dissertations & Theses: A&I*. (Publication No. AAT 9917984).

Rosenshine, B. V. (1981). How time is spent in elementary classrooms. *Journal of Classroom Interaction, 17*(1), 16-25.

Ross, S. L. (1975). An objective grading system for band-orchestra directors. *The Instrumentalist, 30*(2), 28.

Rudman, H. E., Kelly, J. L., Wanous, D. S., Mehrens, W. A., Clark, C. M., & Porter, A. C. (1980). *Integrating assessment with instruction: A review (1922-1988).* East Lansing, MI: Michigan State University, College of Education.

Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education, 58*(1), 37-54. doi:10.1177/0022429409360062

Russo, M. (1988). Written evaluations. *Music Educators Journal, 75*(2), 38-39. doi:10.1177/002743218807500201

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62*(9), 631-634.

Salvador, K. (2011). Individualizing elementary general music instruction: Case studies of assessment and differentiation. (Doctoral Dissertation: Michigan State University). *Dissertations Abstracts International.*

Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education, 38*(3), 57-63. doi:10.1177/002248718703800312

Schleuter, S. (1984). *A sound approach to teaching instrumentalists*. Kent, OH: Kent State University Press.

Schmid, W. (1996). Standards in America's music classrooms (1994-1996). In MENC (Ed.), *Aiming for excellence: The impact of the standards movement on music education* (pp.25-31). Reston, VA: MENC.

Schultz, R. A. (2002). Apples, oranges, and assessment. *Arts Education Policy Review, 103*(3), 11-16.

Schwartz, B. (1996). Communicating student learning in the visual arts. In T. R. Guskey (Ed.), *ASCD yearbook 1996: Communicating student learning* (pp. 79-89). Alexandria, VA: Association for Supervision and Curriculum Development.

Sears, M. F. (2002). A survey of middle school instrumental music assessment methods in southeastern Massachusetts. *Dissertations Abstracts International, 40*(06), 1353.

Seashore, C. E., Lewis, L., & Saetveit, J. G. (1960). *Seashore measures of musical talents.* New York, NY: The Psychological Corporation.

Senty, W. D. (1971). An elementary instrumental progress and grading system. *The Instrumentalist, 25*(2), 22-23.

Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 1066-1101). Washington, DC: AERA.

Sherman, C. (2006). A study of current strategies and practices in the assessment of individuals in high school bands. (Ed.D. Dissertation: Teachers College, Columbia University). *Dissertations & Thesis: A&I*. (Publication No. AAT 3237098).

Shih, Ting-Ting. (1997). Curriculum alignment of general music in central Texas: An investigation of the relationship between the essential elements, classroom instruction, and student assessment. (Doctoral Dissertation: The University of Texas at Austin). *Dissertations Abstracts International.*

Shuler, S. C. (1996a). Assessment in general music: An overview. *The Orff Echo, 28*(2), 10-12.

Shuler, S. C. (1996b). The effects of the national standards on assessment (and vice versa). In MENC (Ed.), *Aiming for excellence: The impact of the standards movement on music education* (pp. 81-108). Reston, VA: MENC.

Shuster, C., Lynch, T., & Polson-Lorczak, M. (1996, April). *A study of kindergarten and first grade report cards: What are young children expected to learn?* Paper presented at the annual meeting of the American Educational Research Association, New York. (ERIC Document Preproduction Service No. ED400289).

Silberman, C. (1970). *Crisis in the classroom: The remaking of American education*. New York, NY: Random House.

Simanton, E. G. (2000). Assessment and grading practices among high school band teachers in the United States: A descriptive study (Doctoral Dissertation: The University of North Dakota). *Dissertations Abstracts International, 61*, 3500A.

Simons, G. M. (1974). *Simons measurements of music listening skills.* Chicago, IL: Stoelting.

Slavin, R. E. (2003). *Educational psychology: Theory and practice*. (7th ed.) Boston, MA: Allyn & Bacon.

Slavin, R. E. (2006). *Educational psychology: Theory and practice.* (8th ed.) Boston, MA: Pearson Education.

Spring, J. H. (1998). *Education and the rise of the global economy.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sproull, L., & Zubrow, D. (1981). Standardized testing from the administrative perspective. *Phi Delta Kappan, 62*(9), 928-631.

Starch, D. (1913). Reliability and distribution of grades. *Science, New Series, 38*(983), 630-636. doi:10.1126/science.38.983.630

Starch, D., & Elliott, E. C. (1912). Reliability of the grading of high-school work in English. *The School Review, 20*(7), 442-457.

Starch, D., & Elliott, E. C. (1913). Reliability of grading work in mathematics. *The School Review, 21*(4), 254-259.

Stetz, F., & Beck, M. (1979, April). *Comments from the classroom: Teachers' and students' opinions of achievement tests.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Stewart, L. G., & White, M.A. (1976). Teacher comments, letter grades, and student performance: What do we really know? *Journal of Educational Psychology, 68,* 488-500.

Stiggins, R. J. (1988). Make sure your teachers understand student assessment. *Executive Educators, 10*(8), 24-26.

Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York, NY: McMillan College Publishing - Merrill.

Stiggins, R. J. (1997). *Student-centered classroom assessment* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Stiggins, R. J. (1998) *Classroom assessment for student success*. Annapolis Junction, MD: National Education Association

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23-27. doi:10.1111/j.1745-3992.1999.tb0004.x

Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice, 20(*3), 5-15. doi:10.1111/j.1745.3992.2001.tb00065.x

Stiggins, R. J. (2005). *Student-involved classroom assessment* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Stiggins, R. J. (2008). An introduction to student-involved assessment FOR learning. Upper Saddle River, NJ: Pearson Education.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement 22*(4), 271-286. doi:10.1111/j.1745-3984.1985.tb01064.x

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practice of classroom assessment.* Albany, NY: State University of New York Press.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and practices, 8*(2), 5-14. doi:10.1111/j.1745-3992.1989.tb00315.x

Stumpo, V. M. (1997). 3-tier grading sharpens student assessment. *Education Digest, 63*(4), 51-54.

Strickland, K., & Strickland, J. (1998). *Reflections on assessment: Its purposes, methods, and effects on learning.* Portsmouth, NH: Boynton/Cook Publishers.

Swanwick, K. (1994). *Music knowledge: Intuition, analysis, and music education*. New York, NY: Routledge.

Swanwick, K. (1998). The perils and possibilities of assessment. *Research Studies in Music Education, 10*, 1-11.

Talley, K. E. (2005). An investigation of the frequency, methods, objectives, and applications of assessment in Michigan elementary general music classrooms. (M. Mus. Dissertation: Michigan State University). *Dissertations, & Theses: A&I* (Publication No. AAT 1428983).

Taylor, G. R. (2003). *Informal classroom assessment strategies for teachers*. Lanham, MD: The Scarecrow Press.

TenBrink, T. D., & Cooper, J. M. (2003). *An educator's guide to classroom assessment*. Boston, MA: Houghton Mifflin.

Terwilliger, J. S. (1971). *Assigning grades to students*. Glenview, IL: Scott Foresman.

Thomas, S., & Oldfather, P. (1997). Intrinsic motivations, literacy, and assessment practices: "That's my grade. That's me." *Educational Psychologist, 32,* 107-123. doi:10.1207/S15326985ep3202_5

Thorndike, R. M. (1997). *Measurement and evaluation*. Upper Saddle River, NJ: Merrill.

Tombari, M., & Borich, G. (1999). Authentic Assessment in the Classroom. Upper Saddle River, NJ: Merrill - Prentice Hall.

Tracy, L. H. (2002). Assessing individual students in the high school choral ensemble: Issues and practices. (Doctoral Dissertation: The Florida State University). *Dissertations Abstracts International, 63*(09), 3143.

Trice, A. D. (2000). *A handbook of classroom assessment*. New York, NY: Longman.

Tuley, R. J. (1985). Framework for performance evaluation. *Music Educators Journal, 7*(9), 32-33.

Tyack, D. B. (1974). *The one best system: A history of American urban education.* Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so had to change? *American Educational Research Journal, 31*(3), 453-479. doi:10.2307/1163222

Walker, D. E. (1998). *Teaching music: Managing the successful music program* (2nd ed.). New York, NY: Schirmer Books.

Walker, G. J. (2001). Assessment: An integral part of instruction. In MENC (Ed.), *Spotlight on assessment on music education* (pp. 87-88). Reston, VA: MENC – The National Association for Music Education.

Watkins, J. G., & Farnum, S. E. (1954). *Watkins/Farnum performance scale.* Winona, MN: Hal Leonard.

Watkins, J. G., & Farnum, S. E. (1962). *Watkins/Farnum performance scale.* Winona, MN: Hal Leonard.

Wells, R. (2001). Curriculum, instruction and assessment: There is a link. In MENC (Ed.), *Spotlight on assessment on music education* (pp. 89-91). Reston, VA: MENC – The National Association for Music Education.

Welter, C. H. (1993). Grade-level assessment in the arts: Of stoppages and stratagems. *Arts Education Policy Review, 94*(5), 2-8.

Wesolowski, B. C. (2012). Understanding and developing rubrics for music performance assessment. *Music Educators Journal, 98*(3), 36-42.

Weymuth, R. W. (1986). The development and evaluation of a cognitive music achievement test to evaluate Missouri high school choral music students. (Doctoral Dissertation: University of Miami). *ProQuest Dissertations & Theses A&I*, 8706495.

Whitcomb, R. (1999). Writing rubrics for the music classroom. *Music Educators Journal, 85*(6), 26-32.

Wiggins, G. (1996). Practicing what we preach in designing authentic assessments. *Educational Leadership, 54*(4), 18-25.

Wilson, R. J. (1990). Classroom processes in evaluating student achievement. *Alberta Journal of Educational Research, 36*, 4-17.

Wing, H. D. (1961). *Wing standardized tests of musical intelligence.* Windsor, England: National Foundation for Educational Research.

Wise, S. L., Lukin, L. E., & Ross, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education, 42*(1), 37-42. doi:10.1177/002248719104200106

Wrinkle, W. L. (1947). *Improving marking and reporting practices in elementary and secondary schools.* New York, NY: Holt, Rinehart, & Winston.

Yeh, J. P. (1978). *Test use in schools.* Washington D.C.: National Institute of Education, U.S. Department of Health, Education, & Welfare.

Zaymeyer, J. L. (1959). Grading with the points system. *The Instrumentalist, 14*(1), 44; 46.

Zdzinski, S. (1996). Parental involvement, selected student attributes, and learning outcomes in instrumental music. *Journal of Research in Music Education, 44*(1), 34-48. doi:10.2307/3345412

APPENDIX A

PILOT STUDY

This is a brief survey relating to how elementary general music teachers in Iowa communicate assessment practices and information to parents at conference times throughout the school year.

1. What grades do you teach? Choose all that apply.
   - ☐ PK
   - ☐ Kindergarten
   - ☐ 1st
   - ☐ 2nd
   - ☐ 3rd
   - ☐ 4th
   - ☐ 5th
   - ☐ 6th

2. How many total years have you taught elementary general music? _____

3. In general, what is the estimated size of your district?
   - ☐ A
   - ☐ 1A
   - ☐ 2A
   - ☐ 3A
   - ☐ 4A

4. How many elementary school buildings are in your district?
   - ☐ 1
   - ☐ 2 to 5
   - ☐ 6 to 10
   - ☐ 11 to 15
   - ☐ 16 or more

   And how many total elementary music teachers?
   (please type an approximate #) _____

5. Does your district have a standard form/format for assessments for parent teacher conferences?
   - ☐ YES - the district mandates one common form for all elementary music teachers to use at conference time
   - ☐ NO - the individual music teachers create their own form for conferences

6. Do you or your district use the same/identical form for each conference reporting period or does the form change each conference period throughout the year?

- ☐ ONE FORM - it stays the same all year
- ☐ MULTIPLE FORMS - throughout the year

7. Which of the following elements does you conference reporting form include? Choose all that apply.

- ☐ concepts covered in class
- ☐ detailed explanation of the concepts
- ☐ indication of individual student achievement on each concept (ie- rating scale)
- ☐ behavioral expectations of all students
- ☐ assessment of individual student behaviors
- ☐ comments - written for each individual student
- ☐ other:

8. Does your form at any time include information that is not taught or covered by the conference reporting time - and is then left partially blank when reported to parents?

- ☐ YES
- ☐ NO

9. What do you find as the most challenging aspect in reporting student progress in music to parents?

10. How do the following topics/areas relate to or impact your assessment forms for reporting to parents?

National Standards in Music Education:

District Standards / Benchmarks:

Iow Core Curriculum:

21st Century Skills:

APPENDIX B

PRE-NOTICE LETTER

November #, 2012

Dear Iowa Music Educator,

A few days from now you will receive a request to complete a survey for an important research project being conducted in elementary music assessment.

The survey focuses on the assessment practices of elementary general music teachers in Iowa, and what factors can impact assessment.

I am writing in advance because I want you to know ahead of time that you will be receiving this survey shortly. This study is important as it could reveal tendencies among music teachers in Iowa and possibly indicate areas for improvement statewide.

Thank you for your time and consideration. It is only with the generous help of music teachers like you that this research will be successful.

Sincerely,

Michelle Hyde Swanson
UNI Music Education

P.S. All of the music teachers who complete and return the survey will be eligible to randomly win a $50.00 gift card to West Music Company!

APPENDIX C

COVER LETTER

December #, 2012

Dear Iowa Music Educator,

I am conducting a study on the **assessment practices and preferences of elementary general music educators** in Iowa. Having taught elementary general music for over 20 years, assessment has always been intriguing to me. Through this research study, I am attempting to discover and describe what is happening in elementary general music classrooms with regard to assessment. I am organizing the assessment research categories into four areas: gathering, organizing, summarizing, and reporting. My goal is to collect accurate information with as many responses as possible to provide enough data to develop an understanding of what is occurring in music assessment throughout Iowa.

As an Elementary General Music Teacher in Iowa, I am asking you to participate in this important study by completing a survey online. The answers you provide will be kept confidential to the extent permitted by law. Special precautions have been established to protect the confidentiality of your responses. You will be asked to provide your name and email address. This information, however, is electronically separate from the survey questionnaire and cannot be connected in any way to any survey responses. Your name is collected only for means of for further correspondence after the study and to exclude you from additional reminders to complete the survey. Your confidentiality will be maintained to the degree permitted by the technology used. Specifically, no guarantees can be made regarding the interception of data sent via the Internet by any third parties.

I would appreciate it greatly if you would take about 20 minutes to respond to the online questionnaire. Your responses, together with others, will be combined and used for statistical summaries only.

The survey is online at: **http://www.take.swanson's.survey.please.com**

Please respond to the questions by ***November 1, 2012***, or as soon as possible. If you prefer a hard, paper copy, please contact me and I will send one to you.

An ***optional*** part of responding to this research survey is submitting any written documents relating to your assessment practices. Please feel free to send electronically (email to: michelle.swanson@uni.edu) any documents that you may use in your assessment tasks (Example: report cards, check lists, communications, grading scales, rubrics, etc.). I can send you a stamped envelope for any documents that you may not have in electronic format.

Each person who completes and submits the survey prior to December 1, 2012, will be eligible to win $50.00 gift card to West Music Company, chosen at random from all of the respondents.

Your participation in this study is voluntary and you may refuse to answer any question. Completing the survey will be considered as your consent to participate. There are no foreseeable risks to you as a participant in this project; nor are there any direct benefits. However, your participation is extremely valued. Following your participation, you may also request a copy of the findings and conclusions, if you so desire. Following research completion, results from the survey will be shared with all participants who provide their name and email address.

If you have any questions about the survey, please contact me at (319) 273-2600 or by e-mail at *michelle.swanson@uni.edu*. If I am not available when you call, please leave a message and I will call back. If you have questions about your rights as a participant in this research project, please contact the University of Northern Iowa Institutional Review Board (IRB) Director of Research Services at (319) 273-6148.

Thank you for your help. I appreciate your cooperation.

Sincerely,

Michelle Hyde Swanson

APPENDIX D

FIRST REMINDER LETTER

November #, 2012

Dear Iowa Music Educator,

About three weeks ago I sent you a request to share your practices in elementary general music assessment. To the best of my knowledge, your survey responses have not yet been submitted.

I truly believe that the results of this research study will be very useful to all music education advocates, and your responses can help get an accurate picture of exactly what is happening in elementary music classrooms in Iowa. Although I sent surveys to people living in every county in the state, it is only by hearing from nearly everyone in the sample that I can be sure that the results are truly representative.

If by some chance you are no longer teaching elementary general music, please let me know and I will check your name off of the list for further communication. Please be assured that individual names can never be connected to the results in any way. Protecting the confidentiality of people's answers is very important to me.

I hope that you will complete the survey as soon as possible online at:

http://www.take.swanson's.survey.please.com

Thank you, again, very much for your time.

Sincerely,

Michelle Hyde Swanson

APPENDIX E

FINAL REMINDER LETTER

December #, 2012

Dear Iowa Music Educator,

I recently sent you a survey regarding the **assessment practices and preferences of elementary general music educators** in Iowa. I am attempting to discover and describe what is happening in elementary general music classrooms with regard to assessment. My goal is to collect information as accurate possible with as many responses as possible to provide enough data to develop an understanding of what is occurring in music assessment throughout Iowa.

I am again asking you to participate in this important study by completing a survey online. The answers you provide will be kept confidential to the extent permitted by law. I would appreciate it greatly if you would take about 20 minutes to respond to the online questionnaire. Your responses, together with others, will be combined and used for statistical summaries only. Please respond to the questions by ***October 14, 2012***, or as soon as possible. If you prefer a hard, paper copy, please let me know and I will send one to you.

The survey is online at: **http://www.take.swanson's.survey.please.com**

An ***optional*** part of responding to this research survey is submitting any written documents relating to your assessment practices. Please feel free to send electronically or any documents that you may use in your assessment tasks (Example: report cards, check lists, communications, grading scales, rubrics, etc.). I will gladly send you a stamped envelope for any documents you wish to submit that are not in electronic format.

Your participation in this study is voluntary and you may refuse to answer any question. Completing the survey will be considered as your consent to participate. There are no foreseeable risks to you as a participant in this project; nor are there any direct benefits. However, your participation is extremely valued.

If you have any questions about the survey, please contact me at (319) 273-2600 or by e-mail at *michelle.swanson@uni.edu*. If I am not available when you call, please leave a message and I will call back. If you have questions about your rights as a participant in this research project, please contact the University of Northern Iowa Institutional Review Board (IRB) Director of Research Services at (319) 273-6148.

Thank you for your help. I appreciate your cooperation.

Sincerely,

Michelle Hyde Swanson

APPENDIX F

SURVEY INSTRUMENT

This is a survey relating to Iowa elementary general music teachers' assessment practices throughout the school year.

This study defines ***assessment*** as the planned and organized use of measurement tools in the classroom where student learning is documented.

I. GATHERING

CONTENT OF ASSESSMENTS:

1. What types of music objectives do you assess in the elementary general music classroom? Please check all that apply.

☐ Singing	☐ Form	☐ Listening
☐ Rhythm	☐ Harmony	☐ Music Appreciation
☐ Melody	☐ Instruments / Timbre	☐ Movement

ASSESSMENT STRATEGIES

2. Please indicate which assessment tools do you use to assess various objectives in your general music classes.

EX = Examinations - written tests to assess students' knowledge of the content area
P = Projects - exhibition that illustrates steps to a final product
A/H = **Assignments & Homework** - activities to practice skills that lead to mastery
O=Observation –individual or group performances in the classroom are observed for assessment purposes
A/V=Audio/Video recordings - observations through audio or video recordings of individual performances
CP = Concert Performance –students are assessed based on their performance in a concert
PF=Portfolios –examples of students' work are kept for the purpose of assessing student growth over time
R=Rubrics –rubrics are presented to students to outline expectations and grading criteria

Objectives:	EX	P	A/H	O	A/V	CP	PF	R	Do not assess
Keeping a steady beat									
Changing tempo identification									
Rhythm identification									
Rhythmic accuracy									
Time signature / beats per measure									
Melodic contour (step / skip)									
Note names (on a staff)									
Melodic accuracy									
Major / minor identification									
Instrument families / identification									
Pitched / unpitched percussion									
Recorder									
Composers / Eras									

Other - specify _____									

Do not assess = none of these listed strategies are utilized to assess this objective

FREQUENCY OF ASSESSMENT:

3. Please mark the term that indicates how often you use the following methods to assess students in your elementary general music classroom. If you never use the method to assess students, mark "never." If you seldom use the method, mark "seldom." If you occasionally use the method to assess students, mark "occasionally." If you frequently use the method to assess students, mark "frequently."

Examinations	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Projects	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Assignments/Homework	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Observation	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Audio/Visual Recordings	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Concert Performances	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Portfolios	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Rubrics	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Other - specify: _____	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently

4. How many* assessments have you administered to each of the following grades since the start of this school year?

** IF you do not teach the given grade level, please select Not Applicable (N/A).*

Kindergarten ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

1st Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

2nd Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

3rd Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

4th Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

5th Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

6th Grade ☐ N/A ☐ 0 ☐ 1-4 ☐ 5-8 ☐ 9-12 ☐ 13-16 ☐ More than 16, specify _____

COMMERCIAL TEST USE

5. Please indicate the grade level with which you used any commercially available tests*.
** IF you do not use the test at any grade level, please select Not Applicable (N/A).*

Musical Aptitude Profile	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Primary Measures of Music Audiation	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Intermediate Measures of Music Audiation	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Music Achievement Test	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Iowa Tests of Music Literacy	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Drake Music Aptitude Tests	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Wing Standardized Test of Music Intelligence	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Measures of Musical Abilities	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Watkins/Farnum Performance Scale	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Farnum Music Tests	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Indiana-Oregon Music Discrimination Tests	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Simons Measurements of Music Listening Skills	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th
Silver Burdett Music Competency Tests	☐ N/A	☐ K	☐ 1st	☐ 2nd	☐ 3rd	☐ 4th	☐ 5th	☐ 6th

II. ORGANIZING

PRACTICES IN ORGANIZING ASSESSMENT DATA

1. Which of the following practices regularly occur within your elementary general music assessment practices? *(check all that apply)*

- ☐ Collecting data for large numbers of students in many class sections
- ☐ Managing data for large numbers of students
- ☐ Identifying areas of strength and weakness from individual student data
- ☐ Analyzing data trends in whole classes or grade levels
- ☐ Transferring observation notes to computer storage

2. Which of the above is your own personal greatest strength(s) when organizing assessment? The following questions are statements to which I seek your agreement or disagreement.

If you **"Strongly Disagree"** with any statement, mark the box next to the **"SD."**
If you **"Disagree,"** but not strongly, with any statement, mark the box next to the **"D."**
If you are "**Unsure"** of your answer, mark the box next to the "**U.**"
If you **"Agree,"** but not strongly, mark the box next to the **"A."**
If you **"Strongly Agree"** with any statement, mark the box next to the **"SA."**

There are no right or wrong answers for these questions. I am only interested in how you feel about the following statements.

I enjoy assessing elementary students in general music.	☐ SD	☐ D	☐ U	☐ A	☐ SA
I find many challenges when completing assessments on my students in elementary general music.	☐ SD	☐ D	☐ U	☐ A	☐ SA
I am limited in my music assessments due to the large numbers of students I teach.	☐ SD	☐ D	☐ U	☐ A	☐ SA
I received adequate training in assessment tools to be used in the elementary music classroom.	☐ SD	☐ D	☐ U	☐ A	☐ SA
I have many professional development opportunities to grow in knowledge of assessment in music education.	☐ SD	☐ D	☐ U	☐ A	☐ SA
I enjoy the measurement issues and terms related to assessment.	☐ SD	☐ D	☐ U	☐ A	☐ SA
There is enough preparation time to assess my elementary general music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
There is enough classroom time with my students to assess all of their skills and knowledge	☐ SD	☐ D	☐ U	☐ A	☐ SA
I could complete more thorough assessment in music class if I had more preparation time outside of the classroom.	☐ SD	☐ D	☐ U	☐ A	☐ SA
The assessments I complete for elementary general music are thorough descriptions of student achievement and growth.	☐ SD	☐ D	☐ U	☐ A	☐ SA

III. SUMMARIZING

ACADEMIC FACTORS

1. The following questions are statement to which I seek your agreement or disagreement with regard to academic factors of music assessment.

If you **"Strongly Disagree"** with any statement, mark the box next to the **"SD."**
If you **"Disagree,"** but not strongly, with any statement, mark the box next to the **"D."**
If you are "**Unsure"** of your answer, mark the box next to the "**U.**"
If you **"Agree,"** but not strongly, mark the box next to the **"A."**
If you **"Strongly Agree"** with any statement, mark the box next to the **"SA."**

There are no right or wrong answers for these questions. I am only interested in how you feel about the following statements.

I consider many academic criteria when factoring student grades such as singing, rhythm, melody, form, harmony, instruments, listening, movement, and appreciation.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Singing is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Rhythm is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Melody is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Form is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Harmony is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Instruments are an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Listening is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Movement is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Music Appreciation is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
If a student tries hard but performs poorly on music objectives, he/she should receive a poor grade.	☐ SD	☐ D	☐ U	☐ A	☐ SA

NONACADEMIC FACTORS

2. The following questions are statements to which I seek your agreement or disagreement with regard to nonacademic factors of music assessment.

If you **"Strongly Disagree"** with any statement, mark the box next to the **"SD."**
If you **"Disagree,"** but not strongly, with any statement, mark the box next to the **"D."**
If you are "**Unsure"** of your answer, mark the box next to the "**U.**"
If you **"Agree,"** but not strongly, mark the box next to the **"A."**
If you **"Strongly Agree"** with any statement, mark the box next to the **"SA."**

There are no right or wrong answers for these questions. I am only interested in how you feel about the following statements.

Participation, attitude, effort, behavior and punctuality are important factors to consider when grading elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Participation is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Attitude is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Effort is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Behavior is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
Punctuality is an important factor to consider when assessing elementary music students.	☐ SD	☐ D	☐ U	☐ A	☐ SA
If a student tries hard but performs poorly on music objectives, he/she should still receive a good grade.	☐ SD	☐ D	☐ U	☐ A	☐ SA

IV. REPORTING

COMMUNICATION TO/FROM PARENTS

1. The following questions are statements to which I seek your agreement or disagreement.

If you **"Strongly Disagree"** with any statement, mark the box next to the **"SD."**
If you **"Disagree,"** but not strongly, with any statement, mark the box next to the **"D."**
If you are "**Unsure"** of your answer, mark the box next to the "**U.**"
If you **"Agree,"** but not strongly, mark the box next to the **"A."**
If you **"Strongly Agree"** with any statement, mark the box next to the **"SA."**

There are no right or wrong answers for these questions. I am only interested in how you feel about the following statements.

I feel that I collect enough data to communicate academic growth accurately.	☐ SD	☐ D	☐ U	☐ A	☐ SA
If a parent questioned a student's grade in my elementary general music class, I could provide documentation to support the grade.	☐ SD	☐ D	☐ U	☐ A	☐ SA

FREQUENCY OF REPORTING / SHARING

2. Please mark the term that indicates how often you use the following methods to assess students in your elementary general music classroom. If you never use the method to assess students, mark "never." If you seldom use the method, mark "seldom." If you occasionally use the method to assess students, mark "occasionally." If you frequently use the method to assess students, mark "frequently."

Report Cards	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Conferences	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Grades	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently
Other - specify: _____	☐ Never	☐ Seldom	☐ Occasionally	☐ Frequently

3. Please indicate the type(s) of report card grades that you assign for your general music classes - checking all that apply.

☐ Letter grades (example: A, B, C, D)
☐ Numeric rating scale (example: 1, 2, 3, 4, and 5)
☐ Descriptive rating scale (example: Excellent, Good, Satisfactory, Needs Improvement)
☐ Plus, Check, or Minus
☐ Emerging, Competent, Mastery
☐ No report card grades for general music
☐ Other - please describe:

CONFERENCES

4. Does your district have set days/evenings scheduled for parent-teacher conferences?

- ☐ YES
- ☐ NO

IF SO, how many times per year? _____

IF SO, as the elementary music teacher, are you... *(check all that apply)*

- ☐ Required to be present at school during conferences
- ☐ Required to be in a certain location at the school during conferences - please list that location: _____
- ☐ Visited by parents to discuss student progress?
- ☐ Working in your office or classroom?
- ☐ NOT required to be present at conferences.

REPORT CARDS

5. Does your district have a standard form/format for music assessment communication to parents?

- ☐ YES - the district mandates one common form for all elementary music teachers to use for communication to parents.
- ☐ NO - the individual music teachers create their own form for reporting
- ☐ NO - I do not include any form/format for assessment communication to parents

6. Does your form/format for assessment communication to parents... *(check all that apply)*

- ☐ Change throughout the year, for the different reporting periods?
  IF SO, what varies?

- ☐ Remain exactly the same for every conference reporting period?
- ☐ Vary for each grade level you teach?
- ☐ Look identical for each grade level you teach?

V. TEACHER OPINIONS

1. What do you believe are educational reasons for using assessments in elementary general music? *Please check all that apply.*

- ☐ Diagnose the strengths and weaknesses of individual students
- ☐ Diagnose the needs of the class as a group
- ☐ Assign grades
- ☐ Evaluate your own instruction and make instructional decisions
- ☐ Communicate academic progress
- ☐ Control students
- ☐ Motivate students
- ☐ Other - please specify:

2. Please indicate any requirements or unique motivators for your classroom assessment *Please check all that apply.*

I believe the following to be important reasons why I assess in my music classroom:

- ☐ establish if students understand concepts
- ☐ monitor student progress
- ☐ allow teachers to adapt instruction
- ☐ motivate students
- ☐ identify gifted students
- ☐ determine students' readiness for next grade level or instrumental ensembles
- ☐ provide teacher accountability
- ☐ assist in assigning student grades
- ☐ provide validity in justifying music program(s)
- ☐ certain number of assessments required by my school district
- ☐ personal reasons
- ☐ Other - please specify:

3. What do you perceive as the greatest <u>challenges</u> related to assessment in elementary general music?

4. What do you see as the greatest <u>advantages</u> to assessment in elementary general music?

VI. GENERAL INFORMATION -

TEACHER DEMOGRAPHICS:

How many total years have you been teaching elementary general music? _____

What is your most recent educational degree? *(circle)* Bachelors Masters Doctorate

What percentage do you currently teach elementary general music?
(circle) Full-time Part-time _____%

How many buildings do you serve? _____

What grades do you teach? Check all that apply:

- ☐ Pre-Kindergarten
- ☐ Kindergarten
- ☐ 1st
- ☐ 2nd
- ☐ 3rd
- ☐ 4th
- ☐ 5th
- ☐ 6th
- ☐ other

How many students do you teach per week (on average)?

- ☐ 100 & under
- ☐ 101-200
- ☐ 201-300
- ☐ 301-400
- ☐ 401-500
- ☐ over 500

What is the average number of students in your elementary general music classes?

- ☐ 10 & under
- ☐ 11-15
- ☐ 16-20
- ☐ 21-25
- ☐ 25-30
- ☐ 31-35
- ☐ 36-40
- ☐ over 40

SCHOOL DEMOGRAPHICS:

In general, what is the estimated size of your school district?

☐ A ☐ 1A ☐ 2A ☐ 3A ☐ 4A

How many total elementary school buildings are in your district? _____

How many total elementary music teachers in your district? _____

Does your school district have a written music curriculum? *(circle)* YES NO

NAME:

EMAIL:

☐ *I would like to receive a summary of the findings of this research study (optional)*

APPENDIX G

FOCUS GROUP PROTOCOL

Focus Group Leader's Statement

You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you made an informed decision about whether or not to participate.

I am interested in any assessment practices utilized in your school for elementary general music. In this group setting, I invite you to discuss the following topics: What are the assessment practices currently being implemented in your music classroom? What are any specific characteristics of the gathering, organizing, summarizing, and reporting practices currently implemented in your music classroom? What are the potential influences on these practices? What challenges or problems do you encounter with assessment in your music teaching? What do you like most about assessment in your music classroom?

I will be video recording our focus group in order to have an accurate transcription. You will be given a copy of the transcript to check for accuracy.

APPENDIX H

INFORMED CONSENT FOR FOCUS GROUPS

UNIVERSITY OF NORTHERN IOWA
HUMAN PARTICIPANTS REVIEW

Project Title: ***Elementary General Music Survey on Assessment in Iowa***

Name of Investigator(s): _____Michelle Swanson_____

Invitation to Participate: You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you made an informed decision about whether or not to participate.

Nature and Purpose: This research study is designed to discover the current practices of elementary general music teachers in Iowa. The current literature reveals no common practices occurring statewide. Two focus group discussions are included as a part of this study to invite local elementary general music teachers to discuss their current practices. The focus groups will follow data collection from surveys distributed to a sample of music teachers across Iowa. The data gathered in each focus group, along with the data from the survey responses, will aid in providing a thorough description of assessment practices of Elementary General Music Teachers in Iowa.

Explanation of Procedures: This procedure will include group interviews, or focus groups, regarding assessment practices in Elementary General Music Education. The purpose of the focus groups is to seek specific practices, characteristics, routines, influences, motivators, inhibitors, or other details relative to assessment in elementary music. Questions will aim specifically at the gathering, organizing, summarizing, and reporting assessment data in elementary music. Each focus group will be video taped and transcribed within one week. Following transcription, each group participant will be sent a copy of the transcription for any comments or corrections. Follow up telephone calls will be made when necessary to assure accuracy. The data gathered in each focus group will be summarized by the researcher

Discomfort and Risks: There are no foreseeable risks to participation. The time commitment for each focus group will likely span from 30 minutes to one hour.

Benefits and Compensation: A potential benefit to participation in this study may include a written summary of all data gathered and conclusions made in this research study. These results will be available to any research participant upon request. Compensation only includes light refreshments during the focus group session.

Confidentiality: Any information obtained during this study that could identify you will be kept confidential by the researcher. In a focus group setting, however, it is impossible for the researcher to promise that other participants will not repeat comments outside the group. Please be honest and open, but please also remain mindful of this limit on your confidentiality. The summarized findings with no identifying information may be shared with participants, and may be published in an academic journal or presented at a scholarly conference.

Right to Refuse or Withdraw: Your participation is completely voluntary.

Questions: If you have questions about the study you may contact Michelle Swanson at 319-273-2600. You can also contact the office of the IRB Administrator, University of Northern Iowa, at 319-273-6148, for answers to questions about rights of research participants and the participant review process.

Agreement: Include the following statement:

I am fully aware of the nature and extent of my participation in this project as stated above and the possible risks arising from it. I hereby agree to participate in this project. I acknowledge that I have received a copy of this consent statement. I am 18 years of age or older.

_____ _____

(Signature of participant) **(Date)**

_____

(Printed name of participant)

_____ _____

(Signature of investigator) **(Date)**

_____ _____

(Signature of instructor/advisor) **(Date)**

[NOTE THAT ONE COPY OF THE ENTIRE CONSENT DOCUMENT (NOT JUST THE AGREEMENT STATEMENT) MUST BE RETURNED TO THE PI AND ANOTHER PROVIDED TO THE PARTICIPANT. SIGNED CONSENT FORMS MUST BE MAINTAINED FOR INSPECTION FOR AT LEAST 3 YEARS]

APPENDIX I

DOCUMENT REVIEW CHECKLIST

Name of School: _____

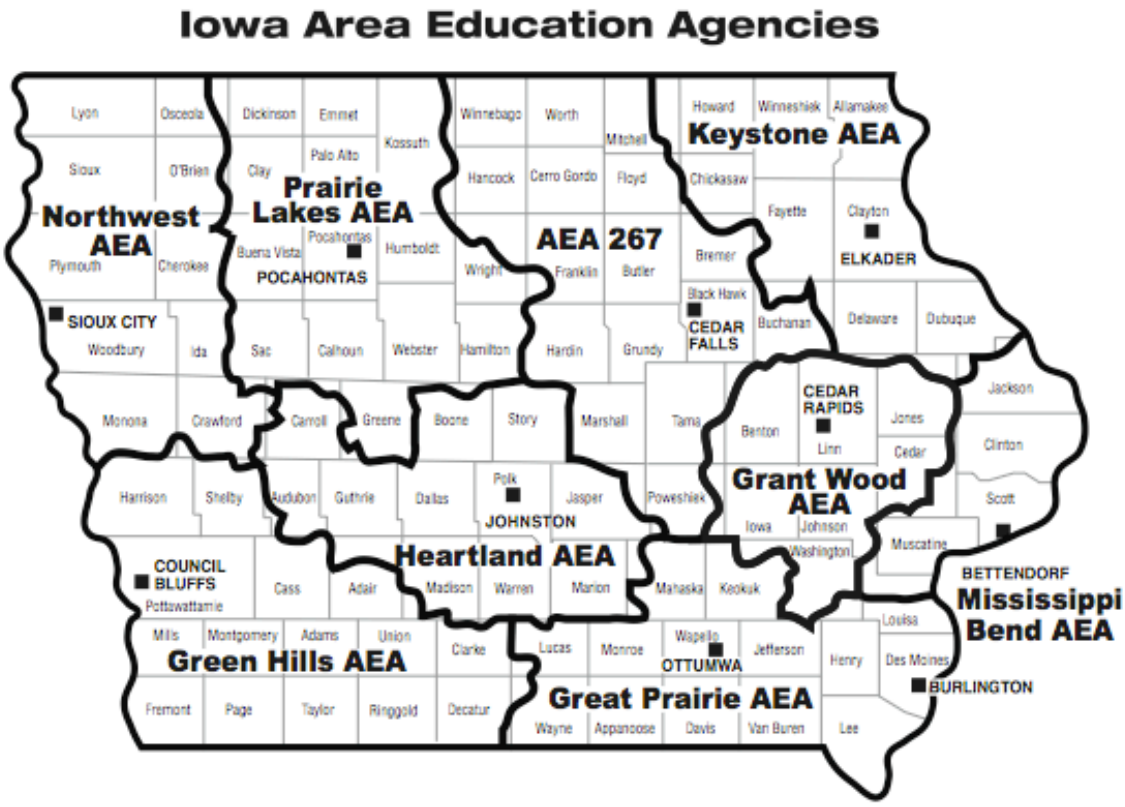
Grade level of Assessment / Report: PKK 1 2 3 4 5 6

Years in use (approximate): 1 2-5 6-10 11-15 16-20

DOCUMENT	YES *or* NO	INCLUDED *and/or* DEFINED	COMMENTS:
- Content/Concept section	Y N	I D	
- Changes with yearly grading periods	Y N	I D	
- Changes with grade	Y N	I D	
- Nonacademic criterion included	Y N	I D	LIST of non academic criterion: Factored into final grade: Y N Separate nonacademic grade: Y N
- Measurement Scale(s)	Y N	I D	TYPE(s):
- Comments shared	Y N	I D	
- District goals noted	Y N	I D	
- National Standards noted	Y N	I D	
- Other general information	X	X	

APPENDIX J

MAP OF AREA EDUCATION AGENCY REGIONS



APPENDIX K

PHONE INVITATION TO PARTICIPATE IN FOCUS GROUP: SMALL DISTRICTS

Researcher's statement:

Good evening. This is Michelle Swanson calling.

I am conducting a study on the **assessment practices and perceptions of elementary general music educators** in Iowa. Having taught elementary general music for over 20 years, assessment has always been intriguing to me.

Through this research study, I am attempting to discover and describe what is happening in elementary general music classrooms with regard to assessment. I am organizing the assessment research categories into four areas: gathering, organizing, summarizing, and reporting. My goal is to collect information as accurate possible with as many responses as possible to provide enough data to develop an understanding of what is occurring in music assessment throughout Iowa.

I have already distributed a survey to hundreds of elementary general music teachers throughout the state. Those results are just being received and I am now following the survey research with two focus groups of music teachers. I would like to invite you to participate in a group discussion about your practices in elementary general music assessment.

This group discussion will be comprised of the elementary general music teachers from school districts similar in size to yours - who are willing to participate. The topic of the group conversation will be the gathering, organizing, summarizing, and reporting practices that occur in your music classroom within your district.

This focus group will occur at Gallagher-Bluedorn Performing Arts Center (in the conference room) on Monday, January # at 4:45 p.m. Light hors d'oeuvres will be served. I anticipate the focus group discussion to be completed in approximately one hour. The consent form will provide many additional details about the opportunity.

May I plan on your participation? If so, I will send you a consent form for you to complete before we begin.

APPENDIX L

PHONE INVITATION TO PARTICIPATE IN FOCUS GROUP: LARGE DISTRICT

Researcher's statement:

Good evening. This is Michelle Swanson calling.

I am conducting a study on the **assessment practices and perceptions of elementary general music educators** in Iowa. Having taught elementary general music for over 20 years, assessment has always been intriguing to me.

Through this research study, I am attempting to discover and describe what is happening in elementary general music classrooms with regard to assessment. I am organizing the assessment research categories into four areas: gathering, organizing, summarizing, and reporting. My goal is to collect information as accurate possible with as many responses as possible to provide enough data to develop an understanding of what is occurring in music assessment throughout Iowa.

I have already distributed a survey to hundreds of elementary general music teachers throughout the state. Those results are just being received and I am now following the survey research with two focus groups of music teachers. I would like to invite you to participate in a group discussion about your practices in elementary general music assessment.

This group discussion will be comprised of the elementary general music teachers from your school district who are willing to participate. The topic of the group conversation will be the gathering, organizing, summarizing, and reporting practices that occur in your music classroom within your district.

This focus group will occur at Gallagher-Bluedorn Performing Arts Center (in the conference room) on Monday, January # at 4:45 p.m. Light hors d'oeuvres will be served. I anticipate the focus group discussion to be completed in approximately one hour. The consent form will provide many additional details about the opportunity.

May I plan on your participation? If so, I will send you a consent form for you to complete before we begin.

APPENDIX M
SMALL SCHOOL FOCUS GROUP: CODING OF THEMES

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Gather	A	Observation and sometimes writing down	Seating chart Boxes Yes / no	Assign boxes (steady beat)	1-3
Gather	B	Observation and write down	seating chart Box Plus - minus (rating scale)	Make sure they get individual attention if needed	8-9
Gather	B	Make a game out of assessing	Singing game Passing game	They don't even know they are getting assessed	15-16
Gather	A	Singing or rhythm Games	Manipulatives (rhythms)		19-20
Summarize	B	So many students	Con		15
Gather	A	Not enough time	quick note	Try to write down observations Can't remember if don't write it during class	24-30
Gather	B	Students unaware	Singing game	They don't know you are assessing them	31-38
Gather	B	Test	Older students	Students know what information will be covered	43-44
Belief	A	Working with others	Small School Only music teacher	Working with another music teacher who is the only elementary music teacher	45-46

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Organize	A	Portfolio	Folder Student Progress	See that they are (or are not) making progress	51-54
Behavior	B	Participation	Points	5 daily points for participation Meeting standards	58-65
Gather	B	Rubric	Participation	Go over it with students	65-67
Gather	B	Rubric	With units	Use with 3rd, 4th and 5th	68-69
Gather	A	Rubric	Want more	Want to develop some more	70
Behavior	B	Attendance at concerts	Communicated to parents	Attendance is a "big part of their assessment piece for their grade"	82-83
Gather	A	Thumbs up / down	Honesty	Sometimes have them close their eyes	100-101
Behavior	A	Participation	Daily Participation Grade		113
Report	A	Standards	Sings, performs, skills, behavior, reads music	Marks standards / adds comments for justification	116-118
Belief	B	Student Accountability	Use for assessment	For what they do	188
Belief	B	Teacher Accountability	Use for assessment	Can adjust teaching based on assessment results	188-192

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Belief	A	Revealing	Confirms or refutes student mastery Use for assessment	Reveals everything about your class	194-197
Belief	B	Revealing	Each class unique	Can adjust teaching based on group assessments	198-200
Report	A	Report Cards Online		End of each trimester	212,
Report	B	Report Cards Online			213
Report	A	Report Cards Online	Pull down menu Rating Scale (+, -)	Share with art and PE	216-221
Report	A	Adding Comments	Online format	Can add additional concerns in writing	227-235
Behavior	A	Effort	Balance of effort and skill	If students are “trying” should they get a good grade?	236-247
Behavior	B	Respect	Behavior categories on report card	Shouldn’t impact other skills assessment	248-250
Report	B	Comments	Online format	“Makes up for lack of clarity on these cards”	280-282
Organize	B	Large Number of Students	373 total students		283

Coding Category		Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Organize	A	Large Number of Students	450 total students			284
Belief	A	Other Duties	Band Lessons		Have to fill day to be full time	285-289
Belief	B	Other Duties	26 band lessons and 5 school duties / week		Less than 2 hours of planning per week	290-292

APPENDIX N

LARGE SCHOOL FOCUS GROUP: CODING OF THEMES

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Gather	A	No standards assessments to use in class Individual preference	District-wide	Can share things if choose	1-2
Belief	A & B	Advantage to large district	Bring in expert for consultation	Professional development days Share ideas	6, 10-12
Belief	C	Mentors	New teachers appreciate veteran assistance		15
Belief	C	Buildings vary	Difference between buildings May not be able to replicate assessments		17-19
Report Summarize	A	Computer No other input			25-26
Report Summarize	C	Comment Box	No list of standards		29
Report Summarize	B	Computer	Other areas have list of standards online	Music is different than other subjects	31-33
Gather	A	“On the fly” “Watching” students	Not “a lot” of “formal assessment”	“I could tell”	53-55
Gather	B	Observation	Majority of assessments		56

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Gather	A	Observation	Looking for student behavior	Looking for students who don't fit	57
Gather	B	Listening	Find ways to have students sing Efficient (time)		58-60
Gather Belief	B	Formal Assessments	Too difficult Can't fit in 25 minutes Not enough time		61-62, 68-70
Gather Belief	A	Schedule	Back to Back Classes With Gym Class	Makes things difficult	72-73
Belief	C	District schedule inequity	Building to Building Not the same within District	Frequency and sections	74-75
Gather Belief	C	Lack of Motivation	Difficult to use assessments that are still motivating	Need a "good lesson" that will "allow you to assess at the same time"	79-80
Belief	C	Overwhelming	Too many things to think about Struggle "for me"	Questions like, "is this going to work?"	80-81
Belief	B	Large Number of Students	Grade level = 70 students Not enough time	No time to review large quantity of student work	91-97

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Behavior	B	Misbehaviors create barriers to assessment	Students not participating Students off task Assessments not reliable Question authenticity of assessment, due to bad behaviors	"Am I really giving them this grade based on their actual skill level?" the behavior gets in the way of demonstrating the skill	101-109
Behavior	C	Participation Grade	Based on student's group participation Student needs to be "doing" the skill or task 3-2-1 rating scale	"It's hard to separate behavior from [skill assessment] sometimes"	110-119
Gather Belief	A	Number of students	570 Really hard to even know their names Challenging		120-122
Belief	C	Accommodations Impossible to do written assessment	Road block to assessment IEP	Student need: read to, one-to-one, away from distractions, write for them, etc.	127-133
Belief	A	Physical Accommodations	Difficult to differentiate	Hearing impairments, wheelchairs, cerebral palsy	134-137
Belief Gather	B	Schedule (timing of classes)	No break in between 6 classes in a row	Grade book (for gathering data) is untouched, due to schedule	138-142

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Organize	B	No Data to organize	Due to lack of time to gather Subjective	Mostly prominent subjective ideas about students, no documented data	145-151
Behavior Summarize	A	Formerly separate NOT on new computer system	Effort and Participation	Used to be able to separate skills and effort/participation Not any more	173-177
Belief Summarize	B	Frustration	Typed comments and were not delivered to parents	Found it frustrating “to the point that this year I didn’t make any comments”	178-184
Report	B	No contact with parents	Parents don’t come to music room during conferences	Comments are that much more important	190-194
Behavior	B	Effort vs. skill	Extra effort overrides lack of skill Lack of effort overrides strong skill	Behavior definitely influences grades, regardless of skill level	193-200
Behavior	A	Grading on Behavior	Shouldn’t grade on behavior	You “can include participation but not behavior”	209-210
Behavior	B	Participation	If not participating, then not demonstrating the skill Students need to consistently participate	A teacher cannot “assume” that a student can “do the skills”	214-219

Coding Category	Tchr.	Key Descriptors	Other Key Terms	Other	Line #
Gather	B	Concert performance / attendance	Performance should be assessment Cannot require attendance	Can't assess if the students do not show up	242-246
Gather	A	Recording	Individual student singing		258-259
Gather	A	Difficult Concepts	Hard to assess music	Not obvious right/wrong answers like other subjects	278-280
Gather	C	Multi-dimensional	Knowledge vs. Performance	Not only have to identify, label, or read, but also perform (sing or play)	283-284
Belief	B	Schedule	Need more time	Music (specialist) schedule should be overhauled	293-297
Belief Gather	A	Large number of sections	46 sections per week	Sections also vary each day of week, due to early dismissals	302-305

APPENDIX O

FOCUS GROUP TRANSCRIPT: SMALL DISTRICTS

A: The primary type of assessment that I do is observation and noting it down, either during class or if it's something that I'm just seeing who gets it and who doesn't and there's enough small numbers that I can get that afterwards. I have on my seating chart this year a bunch of little boxes so that I can assign those boxes to a certain thing and it might be an actual assessment of are they keeping a steady beat on the bordun or it could just be a quick yes or no that they seem to get this concept. So that's the main thing. I am developing something and hopefully will have more information later.

B: I am pretty much doing the same thing. I have them on their seating chart, a little box for them, a "plus" if they get it, "plus-minus" if they're in the middle, and "minus" if they are totally clueless. And then I zero in on the clueless and make sure that they're getting the individual attention that they need, whether it's with me or with a leader/friend, then I'll put them in groups that way. I just go off of our scope and sequence and figure out which scales they're supposed to have and when, and develop my lessons based on that. So, that's basically what I do.

You have so many students that I make it kind of a game when I'm assessing them. And they don't really know that I'm assessing them. Like today we did the "Oh, my no more pie" singing game and I'm passing around a ball and listening to each person sing. I'm writing down about how they are doing.

A: I do a similar sort of thing. I turn a lot of things into games, whether it's singing which is a really easy one to do or rhythm like the game where they're doing a rhythm and they listen and they have a stack of rhythms and they put each one into an envelope in the order that they hear it. And then I know visually right away that the last one, is it the right one? That's hard if I don't know where they got off.

What I find most difficult is, especially for the observations, I try to make a quick note if I can, was it with their singing voice they followed the melodic direction but they are singing down here, or was it that they were trying to get into their singing voice and they ended up over-shooting it. So I try to get that down, too, because otherwise I go back later and try to remember, especially if it was a student that I didn't expect to have a problem, I want to figure out what it is that they're missing, whether it's getting their singing voice or if they just talked through the whole thing. Or something like that.

B: Back to that game thing, I was actually doing this with my first grade students today. I was assessing them on their pitch and making sure they matched high and low. We were doing a little game with the "Brown bear, Brown bear book" and I have little manipulatives in each group and passing them around. I had one student today who said, "I want it to sound like this instead of like sol/mi and sol/mi & la," and she did extra la's

again and she was just making up stuff. But it was with the rhythm and it was ingenious. And I said, "let's all sing it like her!" You find out quickly. I make it a game. They don't know you're assessing them.

With the older grade levels, though, I tell them. If we're going to have, for instance, we're going having a note unit right now with my third through fifth graders to get them ready for playing guitar, recorder and all that---they know there is an assessment.

A: Some of them otherwise get goofy and they would just blow it off.

B: Yea, with the younger grades it's more of a game. With the older grade levels, they know this could be on the test and we need to know this.

A: So, that's mostly performance skills. What I'm working on with another music teacher who is the only elementary in her district, is to develop...a trimester assessment for each grade on written skills so knowing, being able to hear and visually recognize a melody and figure out or write down a rhythm or sight read a rhythm or melody-skills that were expected by the end of third grade, and then give them at each trimester for each of the grades. We're working on developing that for our district.

The long-term effect is that you would be able to have a folder of a student's progress, even if they're behind where you wanted them to be, you could see that they're making progress or not and were drastically dropping off. Maybe they were fine in kindergarten and then first and second grade they weren't making progress, so you could see that.

The only drawback that we're working trying to figure out is presenting it during to the students during music class because that's a big time thing, but it's such a big piece of data that I really want to have in the long term for a student to see their progress.

B: With the daily grades, like daily assessment, I have the daily five points for participation in each class. And, we talk about what participation looks like, what it sounds like, everything about that…in the beginning on day one so they know what is expected of them. Even today…I have mass once a week for my kids, so when they are singing in mass and they are part of the mass, they're either meeting the standards or they're not. Today my third graders were not, and we had a little talk about what does this look like, what does this sound like, and then I went through--do you meet the standards if you are talking to another person?--no. And, I read them the Rubric at the end and we talked about it. We're going to make that up, because the majority of them today were just in la-la land. They're not being attentive. So, we go over that.

"YOU MENTIONED A RUBRIC FOR THOSE EXPECTATIONS. DO YOU USE RUBRIC A LOT OR RARELY?"

B: Sometimes…with a composer unit I do, with masses I do all the time, so I guess that's not a sometimes. Mostly third, fourth and fifth grade I use Rubrics for.

A: I use some. I want to develop some more. It gets to some degree also having the time. Having the "+ -" is, whether it's written down or in your head, you know what the expectations are, so maybe at some point in time for purposes of the school being able to see that and getting that all written down for at least the general expectations of what those mean, so that's along the lines of I have a lot of those types of scoring and I know what each of those is, so I do have some use of Rubrics.

B: My mass Rubric is online so parents can see it.

Also for performances, your performances are your tests, your biggest test of the year for each grade level. For the younger grades, they care so much and want to be there so much. But for the older grade levels--when I taught Middle School--150 points--you do not automatically get those 150 points. If there are teachers finding you talking--they have a Rubric, and we go over it. If they are missing because they went to a game, that's not an excused absence. At the bottom of a note that I send home about it is essential that your child is here and a big part of their assessment.

A: In my district, they don't have a zero or anything on what you do, and the Middle School doesn't necessarily have a policy for that, so as I only go up to 6th grade, and the 6th grade are my creative ones, I couldn't just say that I had to go with what's in there, so for at least that assessment they have to come and sing for me individually so for many, that's a motivator to get them there. If they don't want to sing at the concert, they don't want to sing for me individually--that kind of gets them there. It holds them more accountable. I do see the students that didn't make it and I make them come sing, even if they were sick, because they would have to do that assessment in other classes. That's my view on that. You're not being punished; this is your makeup.

B. When we're studying solfege throughout the year when you teach sol/mi and then sol/mi/la and you keep going--I get to a point -- beginning of 3rd grade -- where I expect them to be able to sing a major scale with solfege with hand signs and that's an assessment piece, and we talk about what it's going to look like for your test. I would demonstrate to them a test. I would have them close their eyes and listen to me sing it and I would sing it totally wrong with repeated mistakes, and then I would sing it and nail it, and then they show me 5, 4, 3, 2, 1. They know what's good and what's not.

A: Assessment of thumbs up and thumbs down--sometimes it's hard for kids to be totally honest so I will usually have them cover their eyes or put it right in front of their chest so only I can see it. I can see where I think they're at, and whether I agree with that or not, and they can also tell me, I'm really, really lost here--or whatever it is--whether they're not getting a new recorder figurine or something is not clicking, or I'm really, really lost, I think I get it, I'm good, ready to go on--and I can go around and give individualized

feedback when needed or can address it if I can see a general problem or I might break it into categories--how are you doing on this part, or how is our tone, how is our volume, how is our rhythm, all different categories--so that we can hone in--if we're preparing for a performance or if it's a piece preparing as a large group-- helping them assess and I can also figure out which students to focus on more.

"HOW DO YOU THINK YOUR ASSESSMENT RELATES TO PROGRAM ACCOUNTABILITY/JUSTIFICATION/CURRICULAR IMPORTANCE OR RESPECT? DO YOU THINK THOSE IDEAS OF RATIONALIZING OR JUSTIFYING YOUR MUSIC PROGRAM RELATE AT ALL TO ASSESSMENT OR DATA, OR DO YOU THINK THEY DON'T HAVE ANYTHING TO DO WITH EACH OTHER? "

A: I think so. I feel with my 5th and 6th grade, that's it's an actual grade that seems to have less relevance. I like the standards K-4 is based for report cards. It just seems with grades there is the expectation that I get an A in music, and so when they don't- or I have a daily participation grade, in terms of what that means, and then assessments or assignments to go along, that kind of tends to freak people out. I would just rather have standards based all the way through. But then my K-4 standards are generalized--sings with the class, performs music skills, behavior, and then reads music for the upper grades. They're so general, I try to justify each thing in my comments, and so I give whatever that trimester's focus was, they're working on tons and tons of things and these are my 3 or 4 key assessments that I want to pull out from this and write about it. I take more time than any other special teacher does in my building for comments, but I want to let parents know that this is why they scored a "plus," that it wasn't just that I was being lazy. They did all this stuff. Or that they need improvement because they aren't doing specific things. I like that I keep my charts and I keep that on there. It's not necessary just for the program but justification for skills that we're working on.

B: With justifying what you do to your administrator, I will pull out all my standards, and I will show my principal I assessed the students "here" or I assessed them "here." My principal will be supportive of that, but then when I get to the report card, it's so generalized with K-5 that you can't possibly--you give them a 4, 3, 2, 1. If I would give a 3, I would almost get… "the parent would be mad about this or that."

A: That's how I feel about the 5th & 6th grade grades.

B: I have all of these assessment pieces that I could show them if they came down, but music is considered in my building--I don't know if it's like that in other buildings--as a special and extra --it's not English or math--I feel it doesn't feel like a core subject because it hasn't been presented that way with the Iowa Core Curriculum which I know people are

working on, but that's hard for me because I have had grades that I have put into the district computer that were changed, and you can't really do anything about it.

A: That's one thing I must say in terms of justification for both of my principals. My principals are very supportive of me. I know I'm a good teacher, but when my principal observes me--I don't think--my principal doesn't know enough about music to be able to give me enough feedback. I know I'm a good teacher and I know I will be a great teacher. Every year I'm learning new things. I know I am doing all these things that are so awesome but I don't get anything back. I have the full support of my administration, but they don't always know what's best for the program and how to support it. And I can tell them these things based on my observations and assessments.

A: I guess my biggest thing I can do follow-up stuff in class but I wish with the focus on giving students individualized time for art--the district has been on PLC's [professional learning communities] helping them with English and helping them with math, what about my music students, too? I have students who are falling way behind in music and I can't ever meet one on one and that's what they need sometimes…that is something I feel pulls me back…that the students who aren't catching up…when I change instruction and try to modify it to their needs, it doesn't help enough.

B: I had a teacher who wanted to pull a struggling student out of music today. And I said, he really enjoys music and I see that he really excels in this class and I just don't think that's a good idea to pull him right now. You can find a better time to pull him and that teacher was OK with that. It was how I felt. If you have a kid that's totally…in music you see a different side of everyone…that's the coolest thing about it. You see those students that struggle in core areas, and they just shine in music. That's what people need to think about--you're trying to take that away.

A: But also students who are struggling with literacy

B: We reinforce it…..

A: They're probably also struggling with that part of music in my class, so if you're going to pull them out for that, first of all, they're not getting that extra reading re-enforcement of decoding, following left to right and all those different things, but then they're going to fall further back. I had a teacher who didn't tell me--just held one of my 6th graders…he was gone so many days he had to catch up" …well, he was gone from my class, too. So now he's farther behind in my music class.

A: One thing I want with the trimester assessments is that it gives them a more continual approach like they have with dibbles for math that they have those, every two weeks or however often they have it--those tests. That shows progress, and if I can say that and show the grade level teacher this is specifically where this student is in music and how I

can then relate it with the supporting of their literacy goals or their math goals where they are in music vs. how they are in your class.

B: I just told students the other day--we were singing FACE "face" and every good boy does fine--and they said, "oh, why do we have to sing that song?" I said, you're going to thank me some day if you're ever in band or if you're ever in music, you're going to need to know these lines and spaces. You will need to know them. Remember the 50 Nifty Untied States? Why do you suppose you can do every state in alphabetical order? Because you put music to it. And when you put music to something, it re-enforces it and you learn it faster. So, if you're struggling in something like reading or math, and you add--make it into--it activates another whole part of your brain.

I've got a student right now, in second grade, who is still struggling in reading….and amazing in music. What I'm doing to help her is I'm making a CD that I found--I'm finding all sorts of different videos on vowels and different sounds and everything is put to music. This girl is going to excel because--the minute I put music in front of her, she gets it. I wish teachers would understand more what we're doing and not just special. I might seem negative, and I'm not--I just want to fight for what I do.

"WHAT DO YOU LIKE BEST ABOUT ASSESSMENT IN GENERAL IN YOUR CLASSROOM? WHAT IS YOUR FAVORITE THING ABOUT ASSESSMENT?"

B: It holds each student accountable for what they do. They hold me accountable. I like that it tells me how I can go home and if kids are not getting something and I've got a bunch of plus/minuses, that they don't get that concept and I need to present it to them in a different way. It's not that they don't get it because they don't get it, it's me. That's my favorite part of assessment--they tell me what I need to improve.

A: Same sort of thing--we touched on this before--getting past in terms of all the different types of assessment--I may think they have something and then the assessment either confirms that or it shows something to the contrary or it shows that I have a pocket of really, really strong students and a pocket of students who don't get it at all, so it reveals everything about your class.

B: And it reveals that not every class is the same in learning. I cannot do the same lesson for both of my kindergarten classes, I can't. One class is totally different---they're just different. I have to tailor it based on what I've seen with the assessment.

A: I've done that before. Last year one of my first grade classes had a really strong sense of steady beat and a really strong sense of what a melodic voice should sound like and the other class just…thumbs down. In terms of everything, as much as I could--and I try to do this a lot anyway--but everything has a steady beat or always feel like there is a steady beat copying me--and then I take my hands away and see how they did, keeping it together as a group and going back to the individual assessment, but were they getting it because they

were copying me or are they getting it because they are looking at their neighbor, and really focused on that with that grade or particular class. And with their singing voices, I'm constantly getting them to sing to reinforce that, because if they don't develop that at that age, then they're just going to fall, fall.

B: And our goal is to build them into musicians and music programs ahead of them.

"YOU SAID YOU HAD SOME REPORT CARDS. DO YOU USE REPORT CARDS AT THE END OF THE TRIMESTER?"

A: They're on line, but I can send you one.

B: Ours are on line, too.

"DO YOU FILL THEM OUT ON-LINE, TOO? NOTHING IS BY HAND?"

A: Ours are on JMC which is our grade program, so I use that with the 5th and 6th grades and their daily grades and their assessments. For them, it's tallying the grade and the comments at the end of the trimester. For my K-4 there's a pull-down menu for each of the four that I said earlier, and it's a "plus," or "blank" is middle, then there's "IN" for improvement noted and a tilde (~) for essentially a minus, or not meeting the standards, and then there's a space for typed comments on the side. Then they get a grade which in elementary is ES, that's plus, that's minus, U and Needs Improvement--all on drop-down menus. Then I submit them. Art, Music and PE are all on one.

B: Ours is very similar to that online. So you have your standards that you have to fill out and then you can put daily grades and enter whatever you do, but for certain grade levels, you can enter all the daily grades you want and they don't affect that final thing at all. If I'm doing all the assessments--if your administration wants you to do all the assessments, and it's not being reflected--what is the point?

A: I go crazy on the comments, just to--I have a ton of things I have written down information for, but I try to pull out the key focus, what they had to know by the end of the trimester, what was just progressive stuff, and try to put those into my comments. To feel that I justify to me--I want to do that anyway--data that doesn't get entered. I like the standards, I wish it could be something where you said this is what they need to know at each trimester and you would have 10 of them, and I could say yes this, no this--all the way down instead of just a general "displays music skills…well they're such a pretty singer," yes, they can sing but they're struggling to read rhythms and they're struggling with the notes are on the staff. They can sing it if I sing it to them.

"DO YOU THINK EFFORT AND ATTITUDE--IF A STUDENT IS NOT MEETING SOME STANDARDS LIKE READING MUSIC OR SINGING ON PITCH, BUT YET

THEY ARE TRYING SO HARD - WHAT DO YOU THINK? OR ON THE OTHER HAND, IF YOU HAVE THE SMARTEST KID EVER, BUT THEIR ATTITUDE IS TERRIBLE, THEY DON'T LIKE TO PARTICIPATE AND AREN'T NICE TO OTHERS--HOW DO THOSE NON-ACADEMIC ISSUES RELATE TO ACADEMIC ISSUES, IN YOUR OPINIONS?"

A: In our current society, again back to the grades, there is that idea that they're in music, art or PE and if they're trying, they should get a good grade. It's a balance because music is not a solo activity. There are professions that can be more solo in terms of what you can do, but in terms of the skills of music, you can do them by yourself, but I tell kids you could be an all-star baseball player but if you don't have teammates, if you're the best pitcher, you can't run and catch the ball in the outfield.

It is a group effort, so if you're not contributing toward the group--if you're not participating--that is part of it, because you're holding back the musicality of the group which is something that I want to assess, so, depending on what is it, are you working to be musical? If you're goofing off, then no, you're not. But then I also think that even if a student tries, and tries, and tries, I also need to assess them on what their actual skills are, so it's more of a balance and leans toward the "is trying" in music.

B: I think mine is a balance, too, on that report card I don't think they influence each other at all. One category on the report card is "respects others" or something like that, so if they're a 1 or a 2, and then they sing great, they've got a 4 over here…

A: In bench marks you can separate that…

B: Yea, in bench marks you can separate that but I think it balances out in the long run because if they have a 4, and then they have a 3, and then a 3, and then they've got a 2, they're probably going to get a 3, so I don't know--it kind of relates, but kind of doesn't--and it should. I have students just like that…Just like you described.

A: Unfortunately, in 5^{th} and 6^{th} grade behavior and attitude do influence grades because of the culture of the school and the expectations lean more toward participation, and that's one reason I wish I could do benchmarks, because I could just mark--your kid is a total goof-off in class but they can sing and they do rhythms, and I wish I could show they do have the skills, but they're not contributing toward the group. It's a combination. It would be one thing if it was an after-school choir, or working on instrumental skills--those could be more separated, but in music class, it's not.

B: I think we could learn something or two, teaching band here for the first year…I think we can learn a thing or two on how they do progress reports. I brought two here, and it's got total quality, rhythm, pitch, dynamics, articulation…if you had in your standards, you either meet it or you don't, and there's a box to check for all of that--it would be so easy, but it's going to take some leadership to get us that way, I'm sure.

A: I just so wish I could do that on each student and…part of it not, because, there's the time, too. Another small-town music teacher and I were talking about how her district wants her to have a certain number of assessments per bench mark, and she has it more like each standard is what they finally decided on because they want at least three assessments and when you have standards plus bench marks in music, however many of them, you can't get three for each and every student. She finally divided it by standards so…maybe it was reads rhythms, ear training, and whatever else.

B: I love that….

A: But at least having specifics… five of them have to be formal assessments and the rest can be observational, because that's what so much of it is, but I could note that day to day to day. On this day I'm really looking for are you keeping the rhythm of whatever it is and the tone and all that. To me that would make it easier on our part.

B: We spend two weeks on these comments anyway, trying to show--to make up for the lack of clarity on these cards. I spend two weeks on my report cards--three hours each night. I have over 200 comments--all developed over 4 years.

"HOW MANY STUDENTS DO YOU TEACH TOTAL?

B: I have 373.

A: Don't know exact number, close to 450, then I have a handful of band lessons.

Because, I have two sections in one of my grades, so I have two empty periods, and they're half period lessons so four total lessons, which makes me full time. They wondered about trying to take out the traveling time so maybe I should have a study hall class. Maybe I could have an after-school choir or a recess choir or something. That would be more effective use of my time and my skills and my expertise.

B: I'm doing band at two different buildings, playing for mass at two different buildings, I do 5 duties a week, and I do 26 band lessons a week. I have an hour and 45 minutes of planning per week.

"HOW LONG ARE YOUR MUSIC PERIODS?"

B: Music periods are 30 minutes, twice a week.

A: Mine are 40 minutes, twice in 6 days. On a 6-day cycle, I see them twice a cycle. So, two times in just over a week.

– January 10, 2013

APPENDIX P

FOCUS GROUP TRANSCRIPT: LARGE DISTRICT

WHAT CIRCUMSTANCES ARE IN A LARGE DISTRICT WHEN IT COMES TO ASSESSMENT? IS ASSESSMENT IN A LARGE DISTRICT BENEFICIAL?

A: We don't have any standard assessments that we use as a district, so it's pretty much up to what each of us decides to do. I think the advantages are that we can bounce ideas off of each other and talk about how to manage the short periods of time that we have and the large groups that some of us have. Again, we can share ideas for things that work for each of us, but we really don't have district assessments.

B: When we do have opportunities to meet, we bounce ideas off of each other and we hear what some of the other people are doing to give ideas of how we can better be doing it in our own classroom, so it is nice to have that collegiality professional network there to compare and contrast.

A: I think another advantage of being in a large district is that sometimes when we have professional development days; we can actually bring in somebody who specifically is tailored to music. And they sometimes will share assessment ideas with us. We just had a guest consultant not too long ago, and he was saying how this could be used for assessments, so that's another advantage.

C: I echo what they said. As a newer teacher, it's nice to come in and have veteran of teachers. I can say, "Is this a good idea?" and so with us there's some of that advice shared. And then you're able to adapt those ideas that are / aren't going to work within your specific building. So even though we're in the same district, building to building we're different. Some assessments might not quite fit well in one building as compared to another.

ARE SOME OF THE ELEMENTARY SCHOOLS IN YOUR DISTRICT" SCHOOLS IN NEED OF ASSISTANCE" AND OTHERS NOT?

A: Right…and some are like the ELA schools.

C: Some are in their first year, and others are in the last year.

A: And there are some that are not on the list "in need of assistance," so everybody is in different phases.

IN YOUR DISTRICT ARE THE ELEMENTARY MUSIC REPORT CARDS ON A COMPUTER OR ARE THEY ON PAPER FOR EACH STUDENT?

C: Yea, they're on computer.

A: It's really just one box and you assign it a number, 1, 2, 3 or 4. There is not much place for input in any other way.

C: There's a comment box, but nothing says, "these are the skills that 5th graders should do on this first report card." There's nothing that we can say "yes" or "no" to.

B: You see that in the other subject areas. You see Quarter 1--here are the skills they should have--Quarter 2 and 3--but within the specials area, there is just a number and there is no delineation between where they lie with their skills.

A: We are eventually anticipating moving more toward standards-based reporting, but it hasn't begun yet, and I don't know that we could really do it yet because our standards and benchmarks are so out of date. We really need to redo all that. We have had that discussion.

C: Were they done soon after the national standards came out?

A: They were done--I want to say in 2000--2000 is about the last time. We have been kind of waiting because the national standards are supposed to be rolling again out in the spring, at least the first new draft. You have to have your standards in place before you can really talk assessment. The state doesn't really have a state assessment for music or any kind of state curriculum…

C: I have even done some research to see what some other state standards are or just to get ideas on what other states' curriculums are like--what they're covering. And, a lot of other states do use standardized music assessments and I think there are a few out there. I think there's one given to 4th graders that I've seen for 4th and 8th graders.

A: Some states have them in 4th and 8th grade. Florida, I know, was working on performance- based assessments but then their funding ran out, and so they never finished with that. So, I know some of them have them, but across the nation it's all across the board.

WHAT DO YOU LIKE MOST ABOUT ASSESSING YOUR STUDENTS?

A: I don't do a lot of formal assessment. I do a lot of ones on the fly. For instance, I was watching my kindergarteners this afternoon and we had just learned a song that was in A B A form.

And they were supposed to be moving during one section of the song and stopping for the other. And I could tell real quickly by who didn't stop, whether or not they were listening and understood the two parts of the song. That's just a real basic one.

B: I think the large majority of assessments that I do is just observation.

A: Looking for the students who just don't fit or demonstrate the standard.

B: That and finding ways to have individuals do something on their own, singing..."goodbye to Kirstin"…and then they sing back "Goodbye to Mrs. _____"…just a really quick...it might take two seconds for me to be able to find out, can they match pitch yet, can they sing up there, and things like that. But as far as formal assessment, it's too difficult, I find, in my 25-minute period to do that.

Even today I did a listening lesson, one that I had never done before. We were listening to a piece of music and then there was a reflection question. We read through the lyrics together and discussed the lyrics and they were supposed to do the reflection question. We didn't even have time to get to the reflection question, so then, do I take a whole another class period to do the reflection question to do the assessment end of it, or do I just say we did the listening, we talked about it and that's good enough. If I had that extra 10-15 minutes, I might have tagged that with written comments, but that's the timing issue for me is what makes assessment difficult. That and the environment…you need a writing board, you need a pencil, and all that takes time out of your 25 minutes.

A: Many of us have back-to-back classes with gym class, and that can make it difficult, coming from or going to gym.

C: We block schedule. And that's the thing, too. The scheduling is also different from building to building, so the frequency is different among us and the duration...

A: When you're working with kids that come from gym and not carrying anything with them, or they're going to lunch and they're carrying their lunches, there's just a lot of management that is necessary, and we don't really have time to do it.

C: The hard part is finding if those are going to be motivating for kids and finding a good lesson that will allow you to assess at the same time. So, there's "is this going to work?" There are too many things we have to think about sometimes.

With listening pieces, I have found S.Q.U.I.L.T. (super-quiet, uninterrupted listening time), that's become a great routine, so for 5 minutes it might be a video that we're watching and talk about it, it might be something they listened to. You might move to something--it's movement--like for form--it's just a way to assess their listening skills--can they be quiet for 3 minutes and listen to a song. Even that's a struggle for some.

It helps with the transition, so they have their recess before they come to see me or if they have a physical activity time in our building that adds an extra physical thing that they have to do, so first grade--you know, last class of the day, and they have that right before specials, so that [the S.Q.U.I.L.T.] helps calm them down when they come to.

SPECIFIC TO LARGER SCHOOL DISTRICTS, YOU'VE MENTIONED MANAGEMENT AND TRANSITIONS, WHAT OTHER KINDS OF CHALLENGE DO YOU ENCOUNTER WITH REGARD TO ASSESSMENT IN ELEMENTARY GENERAL MUSIC? YOU MENTIONED FINDING TIME TO FINISH…

B: One grade level, is about 70 kids. That reflection would have only been the kids adding a paragraph, and then me reading them, and writing a quick comment for each. That assessment was really more for their own self-reflection.

I just had a recorder unit, with a pre-test and post-test. It was hard. The logistics of that, now that I've done it once, I'll go back and I'll know better how to do it, but the logistics…I wanted them to keep track of their practice time and all this kind of stuff…and it just became...to now be faced with these 70 practice books…it's like… am I really going to sit down and go through every single one of these and check these?…I don't know if there is enough time for me to do that for every single student and every single book. So… they wrote it in their letters, and they did this.

For me the biggest issue with assessment is behavior. When I sit down to do grades, I'm faced with…can this student match pitch…I don't know because he never sings because he's always talking to his neighbor and I can never get him to stop talking. I'm constantly having to manage the behaviors and don't do a true assessment…I don't know if he can keep a steady beat because he's never doing that…he's always messing or fidgeting or whatever. So, to me, when I think of my biggest challenge with assessment, because when I sit down to do grades, I am faced with…am I really giving them this grade based on their actual skill level? Or am I giving it based on a behavior that gets in the way of demonstrating the skill. I can't really say.

C: That's the hard part... if you have those students who are not doing a good group effort...they aren't giving back to their whole class, then they just aren't doing the skill in the group. If they're doing something else, they're not on task and they're just not doing it. I grade where I give a participation grade. But if I'm looking for a specific skill and I write down what they get for that skill…if they're on task and doing what they're supposed to do it is a "3." If you know they're not quite participating, maybe because they don't like to dance, they don't like to sing, they're just not going to do it. So, I just don't see it. Then it would be a "2." It's constant reminders and I don't see any work done, then it is a "1." So you try to align with your class what the daily assessment is…it's hard to separate behavior from that sometimes.

A: Sometimes it's just the sheer numbers of kids that you have. In my building we have close to 570 students, and it's really hard, first of all to know exactly who they are, especially in the kindergarten. You have kids coming and going…it's a challenge.

ARE YOU THE ONLY MUSIC TEACHER IN THE BUILDING?

A: Yes…

A: At two grade levels we have five sections of those grades and four places (special classes) to put them, so I take a class and a quarter in kindergarten and a class and a quarter in 2nd grade. So that five classes can go to four specials at a time.

C: Another road block with assessment is how hard it is meet everybody's IEP accommodations. With the third grade group I had last year, five of those students in the class needed either read to, one to one, or away from distractions, or work written for them, or something else. So, it's really hard for a whole class to do a written form assessment if I have to read every question individually to some, and then if another student is supposed to be allowed somebody to transcribe their answers, then I'm also having to write down what he's telling me and I can't really do that all at the same time.

A: Then you have the kids who have real physical needs, too. The kids that have hearing impairments. Today we had three students show up in wheel chairs. We have a new student that moved in and has cerebral palsy, we had two students who came back and had had surgery on their feet and showed up in wheel chairs.

B: Part of that is the timing during the assessments. I have a grade book, but it sits there unused most of the time. I have 25 minutes and then no break. Like today I taught six classes in a row--no break in between--not a 2-minutes--it's like one comes in, one goes out, so then at the end of the day, you're trying to think back to yourself-- trying to remember--and if you try to do it in the middle, you're jumping—when you're trying to do the behavior issues, there's no easy way to keep a physical gradebook. So if I'm being completely honest, if a parent came to me and said, show me why my student has this grade, I would not be able to. It is very much subjective based on what I see every day when they come to class. And then at the end of the grading period, I'm able to think to myself…OK, now what do I see from this student--there's your grade. It's based on my opinion of what I see them doing every day in class. And unless we did happen to have a formal assessment at some point during that trimester, which if we did it would probably be only one or two at most, that would be the only thing I could show them. I know some teachers do a check mark system in their grade books…they go through…she can do it, he can do it…but that's just not practical, especially with the younger kids when they need constant attention…you need to be actively engaged with them constantly. If you're looking down at a grade book, they're going to…

A: You don't even have enough hands to attend to everything if you're playing an instruments and/or you're signing

B: It's not practical…

C: That's kind of why I started doing mine this year where I would write down if it's not skills, it's participation--were you doing it or not--it kind of helps if you have a little something documented--I've tried the clipboard with the student's picture from their seating chart, writing down every grade...I tried that last year…that didn't go. It was just

like a headache because I would forget to write down what it actually was for the date/activity when we did it...it was a really huge mess. And then when you're trying to get a very large group of kindergarteners to sing something by themselves, and then you have four take a turn to share, and then you've lost the rest of them...because of their short attention spans.

C: So, the "singing alone" part of the standards was kind of tricky.

A: Kids have changed over the course of 36 years of teaching...attention spans are definitely shorter now from what they used to be.

C: I just wonder what it will be like when I get to that point in my career!

B: I look at my 3-year-old and I wonder what he is going to be like…he can't focus on one thing for a second…and I think school...

HOW DO NON-ACADEMIC FACTORS FIGURE IN TO ASSESSMENT PROCESS?

A: We used to have ways to mark on the report card for effort and participation, but those disappeared when they redid the new report card and put it on the computer. And, that was kind of a nice way to do it because you could give a grade specifically on their skills but you could also let the parents know that the kids really were trying or they weren't trying or that their participation wasn't…

B: I used to use the comments a lot…in fact, last year I took the time to write out comments, especially if it was a student with the ability to meet the skill requirements in music class, but is consistently not participating or choosing to be disrespectful, the student could excel faster if…you know, things like that. Then I found out that none of the parents got any of my comments, and I was really frustrated by that, because I had sat down and in fact a lot of the grades I did in the in the first two years the parents never got. There was some issue with the printing, so I found that to be frustrating to the point that this year I didn't make any comments because…it's something to do with that unless the parent goes in to look online, they don't see the comments printed out.

B: When I see the report card from my daughter, there are no comments on there anywhere from anybody, and I'm like…well, it would be nice to at least get a comment about something, rather than just see a bunch of numbers.

B: I guess my way of doing that was by writing a comment, and actually some of my comments would end up being quite lengthy because most parents don't come to speak to you at conferences, so that one chance is your chance...unless it's a student that you're consistently calling home about or sending home notes…things like that. If I have a student who gives a strong effort, I have a tendency to put them in the 3 realm regardless, because they're trying really hard, they're being respectful on a daily basis, they're listening,

they're making eye contact, and I can see that they're trying to learn that. To me, that's half the skill right there--is showing me that.

Just the opposite, if a student, even if they can match pitch and keep a steady beat and all that, but they're constantly talking to neighbors and flitting about the room and doing what they want to do, I'm not going to give them above a 2 because they haven't shown me that they can meet those basic requirements.

C: Just before we did grades, at our building during P.D., they kind of talked to us a little bit about grading and stressing it needs to be skill-based, and so there was that piece, you know, what if the student can read at a level 35, but they just don't like to read--they kind of don't try or whatever but they can read it, so they said well, then, that would be beneath grade level if they're a third grader, but if they're surpassing their level but they just kept at the reading thing, so it's like trying to relate that back to how we do special grades. If they're on grade level, is it fair that just because they aren't showing effort, being disrespectful to you, any kind of procedures you have in the room, but then they can still do the skills, they're on grade level.

A: You aren't supposed to grade on behavior. You can include participation as part of your grade but legally you're not supposed to have behaviors in assessment.

C: It's not meant as a punishment if they're going to talk back to me or whatever, if I give you a 2, it's because of your entire skills or if you're just choosing not to participate in something, you didn't do the work today.

B: Most of those types of kids kind of do that for you... when they aren't participating most of the time anyway, they're not demonstrating the skills, so then technically you can give them a lower grade because they're not demonstrating the skill, though you're not supposed to base the grade on behavior, I can't assume you can do the skills just because you demonstrate it once, unless it's a consistent thing. Most of that kind of behavior is…they do that work for me as far as making that decision.

I think the harder one is when you have a student who really, really tries and you want to give them a better grade but you can't. And then usually I will put that in a comment: "shows extremely great effort but still hasn't been able to develop this skill or still has trouble keeping a steady beat," so they know why they are getting a 2 vs. a 3.

IF YOU CALL A PARENT OR WRITE A SPECIFIC COMMENT OR GRADE WITH A SPECIFIC NUMBER, DO YOU SENSE THAT THE NEXT TIME YOU SEE THAT STUDENT THAT IT MADE A DIFFERENCE? IF YOU WRITE SOMETHING AS STRONG AS "THIS BEHAVIOR NEEDS TO STOP," DO YOU SENSE THAT PARENTS ARE TALKING TO THEIR KIDS? THE PERIOD RIGHT AFTER ASSESSMENTS GO HOME, DO YOU SENSE ANY DIFFERENCE?

A: We have “character checks” but it’s not tied to assessments.

B: For specials, we have our own little slip system where it says if they have a rough day-- an “oops form” that says, “today was a hard day in music,” so I can send those home and ask a parent to sign it and send it back. I can check mark what skills they were having trouble with, staying in their seat, respecting neighbors. I see a response from those kind of things for a day or two, but if they’re true repeat offenders, you usually don’t see much long-range change. Most of those children don’t have a problem with what they get as a grade because they don’t care about it. The parents might, but if the kids don’t care, then it makes no difference. Some parents think only about the core areas, “you got this in reading--you’re doing this in math.” In music, they have a tendency to not care, but think...at least in the important areas are OK.

C: I’ve even heard that from some of the kids…if you’re just not going to do your job in here, you’re just going to get a 1. You don’t do anything. And then when we do talk about the grades and report cards… I don’t care. My parents don’t care if I get a 1 in music…it’s just music. It’s not a lot of kids, so I think for the most part those 1’s, it’s usually where they have issues in other areas, where it’s not just my classroom.

B: I would say our greatest fight with assessment is when we have a performance. Technically it could be our strongest assessment but really we have no way of requiring students to be there because it’s outside of school hours and a lot of families don’t make it a priority to have those students there. You can’t use that as an assessment if they are not there. Many students are not in the concert because parents just truly don’t care.

C: Mine are during the day so they have to be there.

B: It might be the way to go... I’m not sure.

A: Sometimes it’s so hard to get parents there for your concerts during the day.

C: I’m torn. I would rather have a student have the opportunity to perform, so have the concert during the day and then all kids get to have that experience -- even if their parents are not there. Or, would I rather have the concert at night so that parents can come, but then some parents would not bring them back for the night performance. So I think I get an okay attendance, for the most part during the day. Since they started doing during-the-day concerts, I’ve had more and more parents come than at first.

I usually do K-2 concert and then 3-5 concert. I tried a family night once and didn’t hardly have any kids.

ANY OTHER THOUGHTS, COMMENTS, IDEAS…

A: The one thing I do is record the kids' singing at least once a year. I do that and then I keep them because they're electronic recordings.

C: As a class?

A: No, individually. I take the time to do it after the concert and with the younger kids, I specify what song they're to do…the ABC song or Twinkle …and we turn it into a performance. But you have to do it in small chunks of kids recording in each class, because their attention spans are short. You can't do everybody the same day…

B: So they would get labeled by their name….

A: Labeled with their name and grade level and when they leave in 5th grade, then I burn a CD for them, which takes time to do because you have to look at all those.

SO, THAT'S LIKE AN ELECTRONIC PORTFOLIO…

A: I was using cassette tapes. It is kind of cool because you can hear the progression of the kids from kindergarten on up, but that does take a lot of time.

C: Another thing, with notation…what kids can play and what they can write down are two different things vs. like with math they're doing these investigations, so they're teaching math as a way to do your own kind of problem solving, so they are able to ask them to explain their answers, show me how you did it, and then be able to write it down, so it's hard to get that piece of music where, you know, show me and explain to me, in organized and in writing--it's not always going to work in music. You just can't.

B: There's only one way to draw a quarter note…

C: And problem solving…

A: If you've got that whole testing piece that comes into play, it's easy to test math and reading skills. And, it's not so easy to test some of the other things that are important for students to learn, but you can't just report with one test score.

C: And if we went to do some kind of formal standardized test in music, I think it would be really hard because then I would feel like I'm just teaching to the test...

A: And whether a kid can identify a symbol has nothing to do with whether or not they can perform it in context.

C: And I would take being able to perform it over just identifying a symbol.

B: Yea, because if they get to 4th grade and they still cannot remember which is a quarter note and which is a half note, but they can tell me…that's one beat, that's two beats…to me, that's the skill that they really need is to know how where it is, but on a test they would obviously want to know…

C: Yea, you could probably interpret it, but maybe not say this is an 8th note pair…

A: Or a kid who can look at a line of rhythm and perform it on an instrument… but they can't tell you what they're doing….

B: And then the schedule...I would gladly do an overhaul of the whole specialist schedule so I could have 5 minutes in between each class to do that function and preparation…to sit down and take a second to…and get ready for the next class. But my problem is, if I did that in the morning and afternoon, that would add an extra 20 minutes on there and on there, and then the specials can't do recess duty, so….

A: You're exactly right…

B: I think if there was more focus on the time we need to actually do that assessment end of it, rather than the superficial…we need people to do our duty or do this or do that duty, then perhaps we would have more success with doing that piece…

A: In a week, I see 46 sections…because I have classes doubled up…46.

B: With the early out (Wednesday), we're on a 4-week rotation, on week 1 we see a section of each grade, week 2 see another. It was not like that when I first came…quarters, but I disliked that because a class would get ahead of the others.

SO, ONE SECTION YOU SEE THREE TIMES A WEEK?

B: Yes, which eventually you rotate and you end up catching up. But what that does, then, is I'm not just teaching the same classes to both 5th grade classes each day…I'm teaching a different lesson every time. Even though I'm doing the same lesson, they're all in a different spot, so I'm not prepping for my day…this is 5th, 4th and 3rd…I have these 6 different classes to teach this morning and these…

A: I kept the Monday & Tuesday and Thursday & Friday to schedule, and then I picked something I could do on those extra Wednesdays that didn't affect the other lesson planning that week--it could be something supplemental that you could add.

B: That works sometimes but then the problem is…if you're in the middle of something on Monday and you don't get to it and then you see them on Wednesday, and then you're going to try to do something else in between

— January 14, 2013