

**KELLY A. PARKES AND  
FREDERICK BURRACK**

**DEVELOPING  
AND APPLYING  
ASSESSMENTS  
IN THE MUSIC  
CLASSROOM**

**ROUTLEDGE**  


*Developing and Applying Assessments in the Music Classroom* is a practical, hands-on guide full of important information and activities for those who want to use assessment as a process for improving student learning and their own teaching. This must-read book pragmatically demonstrates the connection of assessment to curricular goals, program goals, and planned student outcomes by skillfully guiding the reader through all aspects of assessment, including the often-daunting practice of creating meaningful, functional, and efficient music assessments that result in improved music learning for students. This book encourages assessment literacy and skill building to support career-long interest in understanding how students learn, what they learn, and how best to teach them.

– **Ann C. Clements**, *Director, Center for Pedagogy in Arts and Design, The Pennsylvania State University*

There is a void in the literature for music teacher educators who help undergraduate music education students examine how and why assessment is used in music classrooms. This workbook-style text – with learning experiences, worksheets, and activities – certainly helps fill that void. This book is scholarly, practical, and easy-to-read – a must for collegiate music methods classes.

– **Glenn E. Nierman**, *Past President, National Association for Music Education*

*Developing and Applying Assessments in the Music Classroom* is a great resource for developing assessments for specific learning targets, equipping future educators with the tools they need to inspire their students to reach their best potential.

– **Denese Odegaard**, *President, National Association for Music Education*

The learning experiences contained here – each stemming from actual practice and supported by years of research – will allow future and current music educators to develop proficiencies that will be applicable on their first day of teaching and throughout their careers. Parkes and Burrack have carefully delineated the best means to assess each of the inherent facets of music learning: the products of student learning; the processes by which they learned; the programs that are designed to help them learn; and the best practices of the educators who support their learning. *Developing and Applying Assessments in the Music Classroom* will transform the work of music educators and, more importantly, support the vibrant music-learning community that engages teachers and learners in the amazing experiences only music can provide.

– **Douglas C. Orzolek**, *Director of Graduate Programs in Music Education, University of St. Thomas*

With *Developing and Applying Assessments in the Music Classroom*, the editors and authors contribute a valuable resource for future and current music educators at all levels. Through discussions of general principles and music-specific examples, readers will appreciate this book as a companion and guide through the often-confusing landscape of measurement, evaluation, and testing as it is encountered in the teaching of the musical art. It will be equally valuable to individual teachers in their own classroom contexts as it is for administrative leaders of music programs and program-wide teams designing and aligning music curricula. By approaching assessment of musical learning through the perspectives presented in this volume, as music educators we can increase the effectiveness, impact, and authenticity of our educational practice – to the ultimate benefit of our students.

– **Bret P. Smith**, *Associate Professor of Music Education, Central Washington University*



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# Developing and Applying Assessments in the Music Classroom

*Developing and Applying Assessments in the Music Classroom* addresses the challenges faced by today's K-12 educators and future music educators who are expected to utilize and incorporate assessment data as a hallmark of student learning and reflection of effective teaching. Highlighting best practices while presenting current scholarship and literature, this practical workbook-style text provides future music teachers with a framework for integrating assessment processes in the face of a certain lack of understanding and possible dissatisfaction with assessment tools and tasks. Each chapter is prefaced by an overview outlining learning expectations and essential questions and supplemented throughout by an array of pedagogical features:

- Discussion prompts
- Activities and worksheets
- Learning experiences
- Expanded reference lists

Citing examples across a range of musical settings – e.g., band, chorus, orchestra, jazz, and piano and guitar labs – *Developing and Applying Assessments in the Music Classroom* builds from the classroom assessment paradigm, encouraging teachers to create assessment tasks most appropriate to their curricula goals and planned student outcomes. Joined by fellow experts in the field Brian C. Wesolowski and Phillip Payne, the authors invite readers to explore and apply the material in authentic ways to inspire student learning through a comprehensive approach to educative assessment.

**Kelly A. Parkes** is Director and Associate Professor of the Music and Music Education program at Teachers College, Columbia University.

**Frederick Burrack** is Director of Assessment, Professor of Music Education, Graduate Chair for Music, and Distinguished Graduate Faculty at Kansas State University.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# Developing and Applying Assessments in the Music Classroom

Kelly A. Parkes and  
Frederick Burrack

First published 2020  
by Routledge  
52 Vanderbilt Avenue, New York, NY 10017

and by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2020 Taylor & Francis

The rights of Kelly A. Parkes and Frederick Burrack to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging-in-Publication Data*  
A catalog record for this book has been requested

ISBN: 978-0-367-19421-5 (hbk)

ISBN: 978-0-367-19422-2 (pbk)

ISBN: 978-0-429-20230-8 (ebk)

Typeset in Sabon  
by Apex CoVantage, LLC

# Contents

<i>Foreword</i>	ix
RICHARD COLWELL	
<i>Preface</i>	xi
KELLY A. PARKES	
<i>Acknowledgments</i>	xv
<b>1 Historical Foundations</b>	<b>1</b>
KELLY A. PARKES	
<b>2 The Purpose of Assessment</b>	<b>15</b>
FREDERICK BURRACK AND KELLY A. PARKES	
<b>3 Assessment Across a K-12 School Music Program</b>	<b>27</b>
FREDERICK BURRACK AND KELLY A. PARKES	
<b>4 Developing Learning Outcomes, Assessment Tasks, and Scoring Devices</b>	<b>45</b>
BRIAN C. WESOLOWSKI AND PHILLIP PAYNE	
<b>5 Validity, Reliability, and Fairness in Classroom Tests</b>	<b>82</b>
BRIAN C. WESOLOWSKI	
<b>6 Performance Assessments</b>	<b>103</b>
KELLY A. PARKES	
<b>7 Indirect Assessment Techniques</b>	<b>115</b>
FREDERICK BURRACK AND PHILLIP PAYNE	
<b>8 Analyzing Classroom Assessment Data</b>	<b>133</b>
BRIAN C. WESOLOWSKI	

<b>9</b>	<b>Using Assessment for Program Improvement</b>	<b>161</b>
	PHILLIP PAYNE AND FREDERICK BURRACK	
	<i>Glossary</i>	176
	<i>About the Authors</i>	183
	<i>Index</i>	184

# Foreword

Individual assessment was established in music long before Bach auditioned for positions as a church organist. Group assessment in the US occurred in the first few decades of the 20th century with vigorous solo and ensemble contests. At about the same time, psychologists began to wonder whether aptitude in music was anything like IQ tests. Perhaps due to the large influx of students following World War II, college deans became interested in assessment as an entrance requirement for music majors. James Aliferis published an entrance examination (based primarily on music theory skills) that was endorsed by the National Association of Schools of Music. As a fledging public school music teacher at the time, it occurred to me that a college entrance examination and a high school graduation examination for talented students were one and the same; thus began a career-long interest in assessment in music. In the following decade or two there was a flurry of textbooks describing the content of old and new music tests. None of these tests at any level made a significant impact upon the profession and were little used, the primary reason being that music teachers did not know how to administer the tests and little or no idea on how to interpret the results. This book, by experienced specialists, is unique in that it has as its focus the music educator and assessment, preschool through college. The book takes the reader from the establishment of program objectives to their application in the classroom. It is further remarkable in being scaled for use in extant music education methods courses. This means that students will quickly learn the applicability of assessment ideas to their teaching. Despite the rhetoric on arts assessment, the music educator is truly alone. Coursework in assessment given by a college of education or department of psychology seldom applies to teaching and learning in music. School administrators and school board members assume that music is valued and the content rigorous, as statements as early as 1927 by school administrators and later by the National Education Association argue for a “well-balanced school curriculum in which music and the like is to be included with other important subjects such as mathematics, history, and science.” This support

has not translated well in practice. In 1980, John Goodlad reported a strong thread of pessimism among arts educators regarding present and potential support for arts education in the schools. Music, alone, is valued for its sociopolitical goals of education and schooling, but the strength of attaining these goals and others is unknown, requiring a fresh look at assessment. An inspection of the curricular programs of visual arts, dance, and theatre portrays significant differences in their priorities, and there is no documentation of changes in music program objectives to accommodate STEAM or the priorities of process and product. The idea of progress does not hold in the same way as in other subjects. A jazz or rock concert is no more complicated than a performance of Renaissance music. Music performing groups seldom select music in a second year that could demonstrate “annual yearly progress,” an almost infeasible task. The teacher is also responsible for individual and small-group goals (the concert mistress and even the trombones (?) are expected to progress). If the music program is focused on ensembles of iPads or improvising various sounds such as in *Musical Futures*, we know that it is possible to approach any art form naively without any prior conceptions or knowledge of techniques employed and that assessment is even more necessary to avoid the acceptance of curricular issues as experienced by Ronald B. Thomas in the clever *Manhattanville Music Curriculum*. Our assessment practices will be compared with those used in Next Generation Science Standards.

The book was not written as a response to the relentless critiques of the nation’s public schools; its purpose is to enable the music educator to formulate valid and fair tasks and to interpret the results in terms of the student’s and the school’s culture. The authors clearly emphasize the importance of culture, philosophy, and taxonomies in formulating priority objectives that are attainable and measurable. I sense that the user of the materials in this book has the potential to be an instructional leader in music education. This leadership is extremely important in that education itself is searching for leaders and few school administrators are competent to judge valid school music programs. Best practices come from an ability to analyze data and know when data have been interpreted incorrectly, have been manipulated, or are simply false. Only with valid data can a teacher know whether his or her program is rigorous and focusing on values, values that are the basis for teaching objectives in music. Sponsored and collegiate research has traditionally been on small units of music education, not on improving music education programs. That fact, alone, is reason to read this text from cover to cover and become familiar with employing the musical and educational constructs of the profession.

Richard Colwell, Professor Emeritus  
University of Illinois

# Preface

Assessment. This word typically conjures images of tests, feelings of anxiety, and a sense of impending judgment. Public perception of assessment typically centers around either standardized tests or accountability and neither are particularly inviting. The perception of assessment is typically negative. Many, if not most, individuals have had poor experiences with assessment in their own education in K-12 schooling or in higher education. Music educators are not immune to this, and in response to an ongoing need for K-12 teachers to be skilled in generating and using assessment data as part of highly effective teaching, we conceived the idea for this book. We have already published a book detailing a series of assessment tasks and research tools for K-12 teachers (Burrack & Parkes, 2018), yet some research (e.g., Beason, 2017) indicates that K-12 teachers are not aware of the published assessment tasks, nor do they know how to use them. It seems that our higher education colleagues are also, in some instances, not aware there are useful assessment tools that college students could be learning about and utilizing in their coursework/field experiences as they prepare to enter the field.

There is a problem in our profession with what we call “assessment literacy,” and it is perhaps noticeable in higher education faculty at large. Our concern is with those professors teaching in music teacher education programs within universities. Recent research findings (Parkes & Rawlings, 2019) show that some college music education professors report learning about assessment on the job while they were K-12 teachers and others report taking a course more often in graduate school rather than in undergraduate studies; notwithstanding these experiences, many described a certain lack of understanding and highlighted a general dissatisfaction with assessment; and, of more concern, there was little evidence of pedagogical content knowledge used in teaching assessment strategies.

The goal of this book is to provide educative information about assessment to undergraduate music education students in a workbook with relevant research along with learning experiences, worksheets, and activities. Music teacher educators are responsible for preparing future teachers

how to use assessments in K-12 classrooms, yet a series of studies (e.g., LaCognata, 2013; McQuarrie & Sherwin, 2013; Orzolek, 2016; Russell & Austin, 2010) indicate that K-12 music teachers are underprepared to use assessment in their classrooms. Our aim is to provide a text for college music education professors to use within an existing course to educate future K-12 teachers about assessment. We hope this book will serve the need to educate undergraduate music education students about assessment in an approachable manner without adding the need for additional coursework. This book, or selected chapters within, may be used alongside other books in methods courses, practicum experiences, and fieldwork. In providing focused information and activities to develop assessment literacy and skill in undergraduate music education students, we hope to contribute to the education of future music teachers.

The contents of this book are very different to traditional assessment texts that typically deal with classical test theory and test construction. Our book builds from the classroom assessment paradigm: we suggest teachers should create assessment tasks to be used in their instruction that are most appropriate to their curricula goals, their program goals, and their planned student outcomes. We have nine chapters presenting a series of foci, each dealing with different areas of assessment appropriate for K-12 music settings, written in a way that first educates the reader with current scholarship, literature, examples, and information. Each chapter then provides a series of discussion prompts, experiences, or activities and, in some appropriate chapters, assignments for readers to apply the content.

Chapter 1 (Parkes) describes educational assessment and the historical developments that led to an increased focus on assessment in music classrooms. Topics of the chapter include traditional approaches, modern approaches, standardized music tests, classroom assessments, model cornerstone assessments, and technology.

Chapter 2 (Burrack & Parkes) explains the importance of assessment in schools and describes the purpose of assessment as educative. Topics included are assessment processes, recent developments in learning outcomes, the notion of evaluation as part of assessment, and overarching assessment principles.

Chapter 3 (Burrack & Parkes) addresses the rationale for why formal assessment is needed in music programs. The chapter provides a framework for integrating assessment process into music programs. Topics include developing overarching goals, program and course learning outcomes, and detailed instructions for writing outcomes. The alignment of goals and outcomes is underscored and establishing a culture of assessment is recommended.

Chapter 4 (Wesolowski & Payne) describes the processes by which music educators can develop student-learning outcomes, align assessment tasks with stated outcomes, and create scoring devices to collect data in

order to provide detailed, timely, and meaningful feedback and analysis for both students and educators. Topics of the chapter include a review of common educational taxonomies and related frameworks, crafting learning objectives, preplanning assessment processes, developing assessment blueprints and tables of specifications, defining learning outcomes, and selecting appropriate item-types.

Chapter 5 (Wesolowski) addresses the evaluation of music classroom testing quality using three key indicators: validity, reliability, and fairness. Topics of the chapter include defining tests, latent constructs, and inferences; traditional (i.e., large-scale testing) perspectives of validity, reliability, and fairness; and new music classroom perspectives of validity, reliability, and fairness. The new framework for conceptualizing validity, reliability, and fairness in classroom music contexts provides a more accessible, qualitative approach for teachers to reflect upon the quality of their classroom music assessments.

Chapter 6 (Parkes) describes performance assessments and their use in music classrooms. The definitions of direct, authentic, and alternative assessments will be given, and processes for developing performance assessment tasks and scoring devices will be outlined. Topics include developing criteria, performance domains, and the use of performance assessments including portfolios.

Chapter 7 (Burrack & Payne) details strategies to help better understand future students' experience in and throughout a music program. Indirect assessment processes also can expose students' self-concept as a musician and learner of music and beliefs in students' ability to succeed as a musician (musical self-efficacy and musician role identity). These approaches, while often not considered as learning assessment, could be useful to a music teacher to provide proper and effective documentation. Techniques are provided to analyze and compare findings to direct assessments revealing learning influences that can guide instructional and curricular decisions.

Chapter 8 (Wesolowski) describes and analyzes classroom testing data with the intent of informing future teaching and learning processes and improving future test uses from both a class-centered perspective and an individual student-centered perspective. Concepts covered include item- and person-ordering, item difficulty, person ability, item- and person-discrimination, and distractor analyses.

Chapter 9 (Payne & Burrack) addresses the issue of using assessment to enhance program development. Assessment data are collected to guide music teachers toward decisions regarding curriculum, structure, staffing, and student learning through a comprehensive program assessment plan. We have also included a glossary of terms used within the book.

This book has several features that both professors and students may find helpful. Each chapter sets out learning expectations and essential questions at the front, and students will find these useful in developing

their study notes for each chapter. There are *Learning Experiences* placed at certain points in the text that should be completed during the reading. This will allow students to engage with the topic being described and usually these activities are experiential in nature. Activity worksheets are also included at salient points in each chapter to allow students to have hands-on engagement with the knowledge they are acquiring through the text.

Our prime audience for this workbook is college undergraduate music education students and other student professionals completing study to further meet teaching license/certification levels, who want to use assessment as a process for improving student learning and their own teaching. Assessment is simply one part in the cycle of good teaching and learning, and we hope our readers will enjoy learning a new approach to assessment.

Kelly A. Parkes

## References

- Beason, C. (2017). *A mixed methods design study investigating the use of a music authentic performance assessment tool by high school band directors to measure student musical growth* (Unpublished doctoral dissertation). Texas Wesleyan University, Fort Worth.
- Burrack, F., & Parkes, K. A. (Eds.). (2018). *Applying Model Cornerstone Assessments in K – 12 music: A research-supported approach*. Lanham, MD: Rowman & Littlefield.
- LaCognata, J. (2013). Current student assessment practices of high school band directors in the United States. In T. Brophy (Ed.), *Music assessment across cultures and continents* (pp. 109–128). Chicago, IL: GIA Publications.
- McQuarrie, S., & Sherwin, R. (2013). Assessment in music education: Relationships between training, experience, and common practice. In T. Brophy (Ed.), *Music assessment across cultures and continents* (pp. 141–152). Chicago, IL: GIA Publications.
- Orzolek, D. (2016). In their own words: Music teachers' beliefs about assessment. In T. S. Brophy, J. Marlatt, & G. K. Ritcher (Eds.), *Connecting practice, measurement, and evaluation: Selected papers from the 5th international symposium on assessment in music education* (pp. 447–454). Chicago, IL: GIA Publications.
- Parkes, K. A., & Rawlings, J. R. (2019). The preparation of music teacher educators to use and teach assessment. *Contributions to Music Education*, 45, 145–165.
- Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education*, 58, 37–54. doi:10.1177/0022429409360062

# Acknowledgments

We would like to thank Phillip Payne and Brian C. Wesolowski most especially. Without their support, collegiality, wisdom, patience, and hard work, this book simply would not have been written. Not only are they strong musicians and music educators in their own right, they considerably improved the contents of this book with their thoughtful suggestions and powerful contributions. We deeply appreciate their counsel and friendship. We are proud to have them in this book and to count them as our most respected colleagues.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

# 1 Historical Foundations

*Kelly A. Parkes*

## Chapter Overview

This chapter explores the political policies and educational developments that led to an increased focus on assessment in music classrooms in the USA. The chapter explains traditional and modern assessments, with a focus on classroom assessments.

## Learning Expectations for the Chapter

- Describe the political policies that led to modern assessment in music education.
- Describe the difference between traditional and modern assessment measures.
- Identify the difference between standardized and classroom assessments.
- Describe existing measures of music achievement.

## Essential Questions for the Chapter

- What are the differences between traditional and modern assessments?
- What are standardized tests?
- What are classroom assessments?
- How are classroom assessments used as part of educational assessment in music?

## Introduction

When we consider *assessment* at large, it is important to have clarity around the term. Assessment, as explored in this book, is to be considered as **educational assessment**. Assessment in education is a *process* that includes **measurement** and **evaluation**. The assessment process involves

collecting information (data) via a variety of measurement methods that are relevant to making educational decisions. The educational assessment process communicates information about teaching processes and about student learning. This impacts teachers' instructional choices and guides student progress. Measures can be tests, portfolios, checklists, and rubrics. Evaluation occurs when teachers analyze qualitative (words) and quantitative (numerical) data to make a decision of value, worth, merit, or effectiveness. An evaluation is what occurs as a result of an effective assessment process. Evaluations are based on information that has been collected and synthesized to make an educational decision about the extent to which objectives have been achieved. In contrast to evaluations seen on social media platforms (for example, those that determine the quality of restaurants), educational assessments require a wide variety of data and objective measures to guide decisions. Historically, in the USA, learning theories, political decisions, and curriculum reform have impacted educational assessment. Educational assessment is the method by which all aspects of education are measured.

National assessment in the USA started in the late 1960s with the National Assessment of Educational Progress (NAEP, 2016), which included music beginning in the academic year 1971–1972. The NAEP music test includes multiple choice and constructed-response items to illustrate what a nationally representative set of eighth graders know about music. NAEP results represent academic achievement of the nation as a whole. The music portion report (NAEP, 2016) showed that 63% of eighth graders took a music class and that their scores remained similar to the scores of eighth graders who took the test in 2008. These national data suggest that females scored higher on average than their male peers and that students in private schools scored higher on average than those in public schools in music. Although there is a use for large-scale tests, the few that are available in music do not measure achievement of applied music skills or how students apply the knowledge they have attained.

### **Traditional Approaches**

Early approaches to educational assessment were based on objective tests, aligned with behaviorist theories of teaching and learning. Payne (2003) explains that traditional measures were focused on lower-level objectives and structured in a fixed-response style to be machine-scoreable. Early testing would detect correct or incorrect answers on a test by counting the small circles filled in by test takers using Scantron technology. Scantron technology emerged in the 1970s and uses optical mark recognition, identifying dark marks on paper as contrasted by light passing through. Scantron technology is still in use today. These test measures (multiple-choice tests) had evidence of high reliability and were low in cost to administer but also provoked anxiety in students

when used for consequential, high-stakes decisions. These measures are what we refer to as standardized tests, developed out of the psychology field's intelligence testing, where they were adapted to measure classroom content. The origin of this format was introduced in education early in the 20th century to measure student achievement but then quickly expanded to be used to modify curricula and pedagogy. In addition to this, achievement data were also used as screenings for military positions, employment in business firms, and entrance to universities. Student-centered teachers opposed the use of such tests from their inception; however, their use persisted as test developers purported their use to identify teacher competence as well as student learning. The tests did not actually reveal information about teacher competence, yet the notion incorrectly persists today. In the public view, these tests were initially supported because they promised educational accountability, and this has since been cultivated politically as evidence of quality in education (Giordano, 2005). Although mostly outside of music, these standardized tests remain prevalent because they provide a "systematic sample of performance obtained under prescribed conditions, scored to definite rules, and capable of evaluation by reference to normative information" (Payne, 2003, p. 578). They are administered across schools, districts, and states to compare students' performance on the test to other students taking the same test. This is called **norm-referenced** testing. When teachers want to determine whether a student has achieved a particular set of knowledge or a specific skill, within a specific discipline, they would use **criterion-referenced** tests that measure specific criteria related to instructional objectives. In this case, scores of other students are not examined and the results of criterion-referenced tests are usually given as percentages.

Traditional approaches to assessment have revolved largely around multiple-choice and essay formats. When used for numeracy and knowledge attainment, this is reasonable. An essay demonstrates literacy of written communication skills, and multiple-choice tests allow test takers to choose the correct answer to a mathematical problem or demonstrate knowledge retention. When we consider educational outcomes for musicality in performance and decision-making in preparation and the creative processes often expected in music learning, modern approaches offer more appropriate assessment methods. In music, there are defined competencies that result in musical behaviors. For example, (a) accurate reading of printed music allows students to reproduce music in performance or (b) recognizing tonality helps maintain melodic flow in improvisation. We can't identify music reading competency until we hear the outcome during a music performance or, in the case of the second example, recognize understanding of tonality until we hear their melodic improvisation. With numerous competencies being part of music learning, modern music assessment processes are essential.

## **Modern Approaches**

There is evidence indicating that public attitudes toward traditional standardized testing are changing as we see the opt-out movement rising in some states in the USA, in which parents are protesting the ways in which schools assess their children. Pizmony-Levy (2018) reported that 11 states had opt-out rates higher than 5%. New York has seen a 20% opt-out rate from annual tests, Colorado 8% for seventh grade and 11% for eighth grade, and Alaska's opt-out rate is 8.5%. This change in attitude toward standardized testing may come from skepticism about the usefulness of standardized tests as well as the desire for education reform. Educational assessment processes and measures have been developed and improved over the past century due to new theories of learning, curricula reform, and political decisions. Policy decisions such as the Elementary and Secondary Education Act of 1965, the changes made to it in 1994, the No Child Left Behind Act in 2001, and the Every Student Succeeds Act in 2015 have increased the focus on student test scores and the evaluation of teachers and, as such, have impacted how assessment is used in US schools.

Modern assessment processes include classroom-embedded exams, authentic assessments, performance assessments, direct and indirect assessments, and portfolio assessments. Modern measures focus on indicators of learning outcomes to better understand how students understand and apply what has been taught. The purpose of an assessment process is to produce data that can be used to improve learning and guide instructional and curricular improvement. Modern assessment processes focus on objectives that require students to interact with content at a high level of complexity. These processes require of a variety of measures exhibiting student-constructed responses that include oral, written, product-based, and observation formats through which students demonstrate learning. Modern assessment processes can be complex and multidimensional (Payne, 2003), with a high level of validity and high evidence of reliability when developers and scorers are trained. Modern assessments can be less threatening to learners and result in information useful to teachers and their students (Payne, 2003). They take into account cognitive and constructivist understandings of learning, research findings that have resulted from multiple educational studies. Modern assessments are cost-effective measures since most are administered in classrooms. Most importantly, the data can be formatively used to support student learning in the music classroom. These assessments enable teachers to engage with students in classroom activities that combine learning and application, thus constructing knowledge together. The learner gains insight about their progress and the teacher receives guidance on how to plan further instruction to meet the learners' needs. These measures allow teachers to be responsive to learners' needs and, when used alongside

of standardized tests, can give a comprehensive illustration of learners' knowledge and skills.

### *Standardized Music Tests*

Teachers in music have been required to collect student learning data since 2009, after policies were put in place to increase teacher accountability. In 2009 the American Recovery and Reinvestment Act was enacted and provided over \$4 billion to assist educators in meeting the goals of the act. The Race to the Top program (US Department of Education, 2009) was competitive grant funding that rewarded states for developments in reforming education and required some form of accountability. One focus was on developing data systems that measured student growth, ostensibly to also illustrate strengths and weaknesses in teachers. Accountability for K-12 achievement evolved into evaluation systems to illustrate teacher quality by using students' achievement scores. State departments of education required school districts to use large-scale, standardized test data collected in math and reading as evidence of teacher quality. Music teachers (along with other non-tested subject teachers, e.g., physical education) were unable to illustrate their teaching quality because they could not show what their students had learned with data (for a detailed description, see Sherwin & McQuarrie, 2019). Some music teachers were actually evaluated with the math and reading scores of their students, which was a critical misuse of the data. At that time, around 2010, no large-scale music tests existed; however, some of the US states that received Race to the Top funding set about creating them.

A variety of approaches can be seen during the early 21st century in developing large-scale music testing systems as part of evaluating teaching, yet none of them utilized the work of Stufflebeam's (2000) CIPP (Context, Input, Process, and Product) evaluation model to determine quality evaluation systems themselves. This model evaluates educational quality in schools and shows high potential for the evaluation of music teacher evaluation systems (Orzolek, 2019). Research conducted in the early 21st century instead focused on creating measures of music achievement (standardized tests) that would be used, in varying degrees, in teachers' evaluations. Kelly, Cummings, and Gordon (2019) described the Florida Performing Fine Arts Assessment Project that was designed to measure both prepared and on-demand performance tasks, as well as a bank of more traditionally structured test items. Swanson and Shepherd (2019) explained Kentucky's approach, which developed a music listening test that allowed for an authentic application of music knowledge and understanding. Joseph (2019) reported that over five years, Washington state proposed both formative and summative designs to assess music outcomes because formative assessment informs both students and teachers through feedback while shaping instruction, whereas summative