

*Higher Education Music Teacher Educators and Assessment:
Their Understandings, Efficacy, and Satisfaction*

Kelly Parkes & Jared R. Rawlings

*Teachers College, Columbia University * University of Utah*

Abstract

In this study we report what music teacher educators (MTEs, $N = 149$) in higher education understand about assessment. We include their assessment pedagogy, their levels of assessment pedagogy efficacy (APE) at both programmatic (unit level) and personal levels (ProAPE and PeAPE respectively), and the relationship this efficacy has with their (MTEs) satisfaction of assessment pedagogies within their institutions. This mixed-methods study uses a convergent parallel design, with qualitative inductive coding and quantitative factor analyses, correlational analyses, and non-parametric tests. We determine that MTEs report some misunderstanding of the assessment lexicon nevertheless they hold mostly high levels of both personal and programmatic assessment pedagogy efficacy. Differences were observed between MTEs that graduated after 2008 than those who graduated prior to 2008. Findings center on higher education faculty comfort with assessment in higher education with implications for professional development and continued research in the area.

Keywords: efficacy, music teacher educators, pedagogy, satisfaction.

Corresponding author: Kelly Parkes, Music & Music Education, Teachers College, Columbia University, 525 W. 120th Street, #139, New York, NY, 10027.

parkes@tc.columbia.edu

Recommended citation: Parkes, K. & Rawlings, J. (2021) Higher Education Music Teacher educators and assessment: Their understandings, efficacy, and satisfaction. *Journal of Assessment in Higher Education*, 2(1), 35-53. doi:10.32473/jahe.v2i1.122913

©Journal of Assessment in Higher Education. Published under Creative Commons License CC BY- 4.0 International.

Introduction

“Assessment” is a term that has multiple definitions. It can be used to describe one process or several processes that include measures to make an evaluation. Conversely, “an assessment”, or measure, can be used to describe one particular tool, such as a test or a rubric. The focus of this paper is to explore the ways in which Music Teacher Educators (MTEs) in higher education might use these terms. Since MTEs are to prepare future music teachers with skills in assessment of student learning, it is assumed that the terminology is used consistently although there is not confirmation of this. The method and practice of teaching future teachers how to assess uses a relatively new term, Assessment Pedagogy. As a term, it is visible in other subject areas; for example, mathematics pedagogy describes pedagogies involved with teaching mathematics. Assessment pedagogy is not focused so much on the teaching of content but the teaching of how to assess expected learning.

Music Assessment Pedagogy is treated within this manuscript as a discrete area with a separate set of knowledge and pedagogies. This use of the term “assessment pedagogy” can be seen in other disciplines, such as nursing (Croy, 2018), and we suggest that like pedagogical content knowledge (PCK) as described by Cochran, DeRuiter, and King, (1993) and Shulman (1986), there may also be evidence of “pedagogical assessment knowledge” (PAK) (Parkes & Rawlings, 2020) in music teacher educators. In the current study, we explore what MTEs know about assessment and we measure the levels of efficacy MTEs have about assessment pedagogy knowledge. We refer to MTEs’ efficacy perceptions as “assessment pedagogy efficacy” (APE) throughout this study. The purpose of this mixed-methods study was to determine MTEs’ understandings about assessment and to measure their levels of satisfaction and self-efficacy with assessment pedagogy within the context of preservice music teacher education in higher education.

Literature

Efficacy

Efficacy comprises judgments, or beliefs, that an individual holds about their capabilities in certain areas and stems from the motivational work of Bandura (1977). Teacher efficacy includes beliefs about outcomes of student learning and is explained as a teacher’s “judgment of his or her capabilities to bring about desired outcomes of student engagement and learning” (Tschannen-Moran & Woolfolk Hoy, 2001, p. 783). These researchers measured teacher efficacy and reported that a teacher’s efficacy has the potential to influence a teacher’s persistence, patience, enthusiasm, and commitment to teaching. Their study revealed that there are three main constructs within a teacher’s teaching efficacy: instruction, engagement, and management. Assessment efficacy is included as part of the instruction construct, with its own item, in the Teachers Sense of Efficacy Scale (TSES, Tschannen-Moran & Woolfolk Hoy, 2001). It is essential to note that measuring beliefs about teaching, and the constructs within, does not provide

actual evidence of competence, only the self-reported beliefs of the teacher. These beliefs can, however, largely impact a teacher's effectiveness. Teacher efficacy research is growing (Kleinsasser, 2014) and Kleinsasser noted that although the nature of studying teacher efficacy is complex, researchers (such as Hadar & Brody, 2010; Klassen, Chong, Huan, Wong, Kates, & Hannok, 2008; Ruys, Van Keer, & Aelterman, 2011; Takahashi, 2011) have engaged in uncovering useful findings with respect to how teachers' beliefs impact their instructional practice in the classroom along with findings about their levels of satisfaction, as studied by Moè, Pazzaglia, and Ronconi (2010).

Teacher self-efficacy has been studied recently with respect to instrumental and vocal teacher effectiveness (Biasutti & Concina, 2018), pathways into the profession (West & Frey-Clark, 2018), and teacher identity (Wagoner, 2015). While these studies provide some context for the importance of teacher efficacy in music education, they do not specifically target assessment as part of teacher efficacy beliefs. In music higher education, Parkes (2010) reported that applied studio music teachers ($N = 246$) showed generally high levels of teacher efficacy, with the exception of several sub-constructs, such as assessment, that showed lower levels of efficacy. In exploring the data further, Parkes found that 69% of these participants had not received any class, training, or instruction in how to assess student learning in the applied studio setting. This result may be a contributing factor to the ways music teachers learn, or do not learn, about assessment practices and in turn, the beliefs they might have about assessment.

Assessment Beliefs and Understandings

Diaz (2010) argued that knowing how music teacher preparation programs prepare their graduates is important. We suggest that it is also important to know how they prepare graduates to use and teach assessment is also important. We (Parkes & Rawlings, 2019) have previously established that MTEs in higher education are prepared with course content focused on the pedagogy of assessment occurring in graduate study and limited training in the undergraduate programs. In our 2019 study MTEs also reported several concerns with their education with respect to assessment and differences in their understandings became apparent. So, what do these findings mean for music teacher educators and their own assessment pedagogies? The extant literature does not provide answers to this question and there is a documented issue of varied quality when it comes to all teacher-developed classroom assessments (Bonner, 2013). This research points to a problem with how K-12 teachers, in general, are educated in their higher education teacher preparation programs. Fautley (2010) stated that there is a high proportion of disagreement with respect to not only the use of the word assessment, but also how it functions in practice, in classrooms, and how assessment practices are taught. Given the observed confusion around music assessment, and that previous studies illustrate specifically that MTEs have varied exposure to coursework, we suggest that some MTEs may therefore have varied understandings about assessment. They may also have a range of perceptions about assessment pedagogies. We suggest that some individuals may personally feel less confident (in terms of efficacy beliefs) but have more confidence in the assessment pedagogies used by their colleagues at the program or department levels. These beliefs

might be mediated by the length of time they have been in their MTE positions or perhaps the timing of their doctoral training. For example, if they graduated more than 10 years ago, their beliefs may be different than when they joined the academy as junior / pre-tenured faculty in the last 10 years. Therefore, the following research questions guided this study:

1. How do MTEs describe or define assessment and its importance within their coursework?
2. What are the levels of satisfaction, personal assessment pedagogy efficacy (PeAPE) and program assessment pedagogy efficacy (ProAPE) as reported by MTEs?
3. Is there a relationship between the amount of ProAPE MTEs report and their levels of personal assessment pedagogy efficacy PeAPE?
4. Do MTEs graduating after 2008 have higher levels of ProAPE and PeAPE than MTEs graduating prior to 2008?
5. To what extent do the quantitative item results and the qualitative item results align with one another?

Method

Mixed-methods design incorporates qualitative and quantitative methodologies in a complementary approach to understand a research problem (Creswell & Plano-Clark, 2011). We chose a convergent parallel design to develop a more complete understanding of MTEs' beliefs and perceptions of assessment, including its pedagogy, within music education coursework. Convergent parallel design involves qualitative and quantitative data that is collected "concurrently but separately—that is, one does not depend on the results of the other... [and] have equal importance for addressing the study's research questions (Creswell & Plano-Clark, 2011, p. 78). We utilized a "data-validation variant" (Creswell & Plano-Clark, 2011, p. 81) of convergent design and both the quantitative and qualitative strands were collected concurrently.

To determine music teacher educators' knowledge of assessment and their satisfaction and efficacies (both personal and program) with assessment pedagogy, we conducted a national survey of all faculty listed at schools listed with a NASM undergraduate or graduate music education teacher preparation program. A graduate research assistant confirmed, modified, and/or deleted information for all MTEs on the published contact list resulting in 1,500 potential respondents and we invited the entire contact list to participate in the study via email. Overall, we received 149 survey responses with our items of interest completed, indicating a response rate of 9.8%. Following the protocol prescribed by Lohr (2008) for the sample mean, we calculated a +/- 7.6% margin of error (CI: 95%), which allowed us to pursue further data analysis. It should be acknowledged that of the 1,500 solicited, our participants only represent individuals that willingly chose to respond to a research survey about assessment. This may represent only those individuals that are interested in, or have a certain predisposition towards, assessment. The majority of MTEs emailed chose not to respond to our research invitation and the significance of this will be addressed later in the paper.

Participants

In this study, we examined only those participants that completed the questionnaire items with respect to assessment understanding and importance ($N = 149$), along with those who answered all items for satisfaction with assessment and efficacy for assessment pedagogy ($n = 141$). The participants are the same population as those reported in our earlier study (Parkes & Rawlings, 2019). That study focused only on how MTEs are educated about assessment. In both studies, participants varied by their university faculty post (6.0% adjunct / part-time, 17.4% tenure-track/non-tenured, 22.8% career line / full-time, 53.7% tenure-track / tenured). For a full description, please see Table 1.

Table 1

Description of Participants

Gender	n
Male	69
Female	77
Non-conforming	2
Non-response	1
Ethnicity/ race	
White	131
Asian	1
Black/ African-American	2
American-Indian or Alaska native	2
Hispanic, Latinx, or Spanish	1
Other	4
Non-response	8
Level of education	
Ph.D.	95
Ed.D.	18
D.M.A or D.Mus.	14
M.M.Ed or M.Mus.	12
M.A or M.A.Teaching	2
M.Mus.	3
Other	5

*n = 141 participants represented 38 states, with 8 not disclosing their location.

Survey Instrument Development

Our survey asked participants to estimate and describe their perceptions of multiple topics about assessment, its practice and pedagogy, through closed- and open-response items directly addressing our research questions. We also asked a series of demographic questions (such as rank, experience, graduation from graduate schooling, state location etc.) To develop the survey, and prior to data collection, we invited two experts, experienced MTEs not located within NASM-accredited institutions, to take the complete survey and to participate in a 20-minute interview session about the design of the survey. During these cognitive interviews (for details of this technique and rationale for use, see Desimone & LeFloch, 2004), we asked the two MTEs questions about the clarity of the survey and they reported it took between 20–25 minutes to complete. They suggested modifications to response options and other corrections. This protocol ensured that the final survey had some evidence of face validity. The data sets from these two survey responses were destroyed and not used as a part of the data file for the study.

For the present study, we used three items to explore what MTEs understand about the term assessment and how important assessment knowledge and practices are in the preparation of teachers. (Please see Appendices in supplemental material for all items). We used another 10 items specifically about satisfaction and efficacy beliefs for assessment pedagogy. These 10 items focused on measuring individuals’ basic level of satisfaction with assessment pedagogy in their respective program units, along with measuring their personal and program levels of efficacy with respect to assessment pedagogy. See Table 2 for item means. Items 1- 4 in each section were adapted from Tschannen-Moran and Woolfolk Hoy’s (2001) measuring teaching efficacy, which included an item focused on assessment.

Table 2

Item means

Items	Mean	SD
PeAPE		
Authentic assessments	4.602	.7820
Teaching assessment	4.588	.7146
Providing multiple examples	4.801	1.024
Assist student teachers in K-12	4.477	1.128
Knowledge of assessment standards	5.117	.958
ProAPE		
Authentic assessments	4.1418	.9225
Teaching assessment	4.5035	.9533
Providing multiple examples	4.4255	.9579
Assist student teachers in K-12	4.4184	1.0152

*Cronbach Alpha for all 9 items = .865

Analyses

Merging quantitative statistical results with qualitative findings offers an opportunity to further substantiate and explain relationships. Quantitative data were collected via the adapted efficacy instrument and qualitative data were collected via open-ended questions asking participants to share their beliefs and perceptions about assessment. For research question 1, we employed qualitative analyses, and for research questions 2-4, we used quantitative analyses. To address our mixed-methods research question, we utilized a strategy to merge the two sets of results. Creswell and Plano-Clark (2011) suggest “identify[ing] content areas represented in both data sets and compare, contrast, and/or synthesize the results in a discussion or table” (p. 79).

Qualitative

The research questions for the current study required computer assisted qualitative data analysis software (CAQDAS) to organize the large qualitative data file. Participant data were imported from the survey platform Qualtrics into NVivo 11.3. As a way of ensuring trustworthiness to our analyses, we structured the coding process into two stages: first and second cycle coding (Miles, Huberman, & Saldaña, 2014). Author 2 used NVivo to assist with the coding of these data and Author 1 used traditional methods of open coding. We explored the data file using a holistic method of coding (Miles et al., 2014) and independently agreed on the preliminary codes by utilizing a focus prompt approach (Kane & Trochim, 2007). During the second cycle of coding, we selected a pattern coding protocol to condense, synthesize, and elaborate the data. Lastly, we used axial coding to group like codes into larger themes (Patton, 2015). We also chose a frequency count analyses for two questions, for appropriateness in summarizing those data.

Quantitative

We conducted descriptive and factor analyses to estimate mean scores, factors, and reliabilities. The researcher-created instrument was designed to measure beliefs about their teaching of assessment and the level to which their program teaches about assessment in the preparation of pre-service music educators. In order to examine evidence of internal consistency (reliability), we calculated Cronbach’s alpha for the 9 items (knowledge of state requirements and 8 efficacy items). We found good evidence of internal consistency ($\alpha = .86$) across the 9 items and considered this appropriate to undertake further analyses.

For these analyses, a data file was created in SPSS 24.0 for Mac by downloading the raw data from Qualtrics and these data were also screened utilizing techniques associated with missing data protocols (Pallant, 2013). Participant descriptive statistics were calculated to explore the data file for normality and outliers. To allow for parsimonious correlational analyses, we computed variables by calculating the mean

scores for each participant's self-reported "personal assessment pedagogy" efficacy (*PeAPE*), "program assessment pedagogy efficacy" (*ProAPE*), and satisfaction with program assessment in general. We then conducted Spearman's Rank Order Correlation (ρ) to determine relationships between assessment efficacy scores (personal and program). Response options for reporting graduation year were recoded into a dichotomized option for MTEs graduating prior to 2008 and those MTEs graduating in 2008, to divide the sample evenly. Finally, we calculated a series of non-parametric tests to assess differences between these two groups of participants, Mann-Whitney U Tests to determine levels of difference in *PeAPE* and *ProAPE* between graduation years.

Results

Research Question 1A: How do MTEs describe or define assessment and its importance within music education coursework?

Music teacher educators described and defined assessment, as a word and as an enacted part of teaching, as a practice. We present evidence from the MTEs in a condensed narrative form based on the richness of the aggregated data. We asked participants to describe their understanding of the meaning of the word "assessment." Most participants (68%) described "assessment" writ large, as a *process* or series of *processes*, while 32% also described the "assessment" as a series of *data collection tools* or as *product-oriented*. In the following segment of the paper, we list each theme and the sub-themes (processes and products) that informed our coding decision.

Processes

MTEs in this study report assessment being a process related to student learning. One participant wrote that "assessment provides data that helps to determine student understanding and learning and provides feedback on the efficacy of material presented, leading to revision on instruction." Another participant remarked "assessment is a general term that encompasses informal approaches to gathering information about student learning and providing feedback and formal processes of measurement design, data collection, data organization, evaluation, and reporting." Although many participants described assessment with similar words, both quotes directly capture how the MTEs in the current study believe assessment is a process of student learning. For example, "It [assessment] can include tests (systematically gathered materials), evaluations (judgments as to quality), informal or formal work".

Product oriented

This theme emerging from participants' reports describes assessment as a data collection tool and/or a measure of teacher effectiveness. One participant commented that "assessment is the tool or tools by which students demonstrate their understanding, knowledge, or skills on a particular topic" which is perhaps not illustrative of assessment as a process but instead *an* assessment used as a tool.

Approximately one-third of the participants discussed data collection tools as a means of measurement within their description of assessment, whereas a few mentioned a difference between the terms “assessment” and “measurement.” Of the participants that wrote about this, some were adamant that measurement is different to assessment. Also, that they shared that the product is explicitly seen as evaluation “when you apply a grade to the activity”... as “a final activity at the end of class”.

To summarize, while we report briefly here the differences in how MTEs in our study described and defined assessment, we note that compared to the current measurement, evaluation, and assessment literature, many of the descriptions and definitions given by our participants are slightly different to the definitions held in the field as we defined earlier. *Assessment* is a term used to describe a process of measuring and. *An assessment* is used to describe one particular measure, such as a test or a rubric. An assessment is perhaps not seen widely as a product, yet many of our MTE participants stated this was how they viewed it.

Research Question 1B: How important is assessment to MTEs in the preparation of music teachers?

We asked participants this question directly in item number 3 of the qualitative section of the questionnaire. Frequency analyses indicated that 105 participants (70%) responded it was “extremely important”. Another 40 (27%) indicated it was “moderately important” with two (1.5%) responding it was “slightly important” and two (1.5%) responding it was “slightly unimportant”. No participant indicated that assessment was either “mostly unimportant” or “extremely unimportant”.

We also asked participants to describe how they viewed assessment within the context of music teacher education and music teacher preparation, in a parallel item to the importance item, to garner a richer exploration. Most music teacher educators (73%, $n = 107$) answered this open-response question, while 27% did not answer. Of the participants that answered the question, 55% ($n = 58$) wrote about the importance of assessment or that assessment was valuable to music teacher education. Thirty percent ($n = 17$) wrote that assessment enhances teacher practice and K-12 student learning. In the following segment of the paper, we report each theme with evidence from the participants.

Music teacher educators reported that assessment, within the context of music teacher education, is important and valuable. Broadly, many participants believe that assessment is “essential,” “crucial,” or “vital.” One participant wrote, “assessment should be implemented from day one of admittance into the teacher preparation program.” Indeed, another participant remarked, “Assessment is an important part of the teaching/learning process, so I believe it's important for preservice music teachers to know how to assess students' achievement levels...Assessment also helps the teacher to evaluate his/her teaching effectiveness.”

Another participant stated, "Assessment is crucial to music teacher education. It provides the main source of information regarding students' understanding and ability, and should inform planning and instruction. Without assessment training, teachers may not be prepared to grapple with data and evaluation."

Music teacher educators shared that assessment, within the context of music teacher education, should enhance teacher practice and K-12 student learning. One participant observed "Musicians assess constantly! Making assessment conscious and specific makes music teaching and learning more effective and efficient." Likewise another participant shared; "In order to improve teaching practice, music education students should discover or create realistic strategies that will help them gather information from their students that can be used to improve teaching practice." Yet another participant remarked that "Assessment must be included in music teacher education so that our students learn how to know of their own students are learning...Assessment procedures are and will remain essential for any universal education system."

The benefit of comparing both qualitative and quantitative data here allows us to explain in more detail why our participants feel assessment is important. This explanation connects to the data revealing MTE perceptions of efficacy and satisfaction.

Research Question 2: What are the levels of satisfaction, personal assessment pedagogy efficacy (PeAPE), and program assessment pedagogy efficacy (ProAPE) in music teacher educators?

We first calculated personal (PeAPE) and program assessment pedagogy efficacy (ProAPE) scores separately. For mean scores and evidence of reliability, see Table 2. The PeAPE items yielded a separate Cronbach's $\alpha = .751$ and the ProAPE items, $\alpha = .845$; together all nine items yielded Cronbach's $\alpha = .865$.

A factor analysis was performed using the Principal Component method of extraction to determine whether there were identifiable factors within the instrument. Bartlett's test of sphericity, which tests the overall significance of all the correlations within the correlation matrix, was significant $\chi^2(36) = 563.178, p > .000$. The Kaiser-Meyer-Olkin measure of sampling adequacy indicated that the strength of the relationships among variables was high (KMO = .84), thus it was acceptable to proceed with the analysis. We also examined the ratio of participants to variables, which was illustrated with 141 participants and nine items. This yielded a ratio of 15:6 which is above the recommended ratio of 10:1.

The first factor held an Eigen value of 4.51 and it accounted for 50.20% of the variance. The second factor, with an Eigen value of 1.05, accounted for a further 11.72% of the variance. To ascertain the loadings on each item, we used a Promax rotation method with Kaiser Normalization, converging in three iterations. See Table 3 for the pattern matrix.

Table 3

Pattern Matrix

	PeAPE items	ProAPE items
Authentic assessment experiences	.912	
Teaching assessment	.913	
Program Authentic assessment experiences	.730	
Program Teaching assessment	.706	
Multiple examples	.661	
Program multiple examples	.591	
Assisting K-12 students		.901
Program assisting K-12 students		.798
Knowledge of Assessment		.558

*Extraction method: Principal Component Analysis
 Rotation Method: Promax with Kaiser Normalization ^a
^a Rotation converged in 3 iterations.

The two components were correlated ($r = .547$) so we can report that the personal and program efficacy beliefs are one construct; however, items asking about assisting student teachers to conduct assessment for K-12 students (both program and personal) and assessment knowledge are separate. This result seems straightforward as the student teaching setting is typically removed from the coursework usually associated with assessment techniques and is often mediated by the cooperating teacher in the field setting. MTEs have reasonably high levels of both personal (PeAPE) and program assessment pedagogy efficacy (ProAPE), with slightly lower efficacy with respect to teaching student teachers how to develop assessments for K-12 students, at both personal and program levels. MTEs report a very high level of knowledge with respect to state education standards and requirements.

The levels of satisfaction were generally high with most MTEs ($n = 123$; 87.5%) satisfied with program assessment pedagogy. Specifically, 12.5% were extremely satisfied, 56.58% moderately satisfied, 18.42% slightly satisfied. Just over 12% ($n = 17$) were in any way dissatisfied; with 7.89% lightly dissatisfied, 3.29% moderately dissatisfied, and 1.32% extremely dissatisfied.

Research Question 3: Is There A Relationship Between The Amount Of Program Assessment Pedagogy Efficacy (ProAPE) MTEs Report And Their Levels Of Personal Assessment Pedagogy Efficacy (PeAPE)?

The relationship between program-level efficacy with assessment processes and personal levels of efficacy with assessment pedagogy was investigated using Spearman’s Rank Order Correlation (ρ). Preliminary analyses revealed a violation of the assumption of normality. According to Cohen’s (1998) determination of strength of the relationship, there was a strong, positive correlation between the two variables, $r =$

.61, $n = 141$, $p < .01$, with high levels of program-level efficacy associated with high levels of personal self-efficacy with assessment pedagogy. The strength of the correlation, R^2 , was .36, indicating that program-level efficacy helped to explain 36% of the variance in respondents' scores on the personal self-efficacy scale. This can be considered a reasonable amount of variance explained when compared to much research in the social sciences.

Research Question 4: Do MTEs graduating after 2008 have higher levels of ProAPE and PeAPE than MTEs graduating prior to 2008? (This question is separated into three sub questions in order to complete the Mann-Whitney U tests).

Research question 4a: Do MTEs graduating after 2008 have higher levels of personal efficacy than MTEs graduating prior to 2008?

A Mann-Whitney U Test revealed a significant difference in the personal efficacy levels of MTEs. MTEs that graduated after 2008 have significantly lower levels of self-efficacy with assessment pedagogy ($Md = 4.33$, $n = 59$) than MTEs that graduated prior to 2008 ($Md = 5.00$, $n = 77$), $U = 1625$, $z = -2.88$, $p = .004$, $r = .25$.

Research Question 4b: Do MTEs graduating after 2008 have higher levels of program efficacy than MTEs graduating prior to 2008?

A Mann-Whitney U Test revealed a significant difference in the program efficacy levels of MTEs. MTEs that graduated after 2008 have significantly lower levels of program efficacy with assessment pedagogy ($Md = 4.00$, $n = 59$) than MTEs that graduated prior to 2008 ($Md = 4.50$, $n = 78$), $U = 1804$, $z = -2.18$, $p = .03$, $r = .20$.

Research Question 4c: Do MTEs graduating after 2008 have higher levels of satisfaction than MTEs graduating prior to 2008?

A Mann-Whitney U Test revealed a significant difference in the satisfaction levels of MTEs. MTEs that graduated after 2008 have significantly higher levels of satisfaction with assessment pedagogy ($Md = 2.50$, $n = 59$) than MTEs that graduated prior to 2008 ($Md = 2.00$, $n = 77$), $U = 1653$, $z = -3.04$, $p = .002$, $r = .26$.

Quantitative Findings Summary

Quantitative results demonstrate that MTEs perceive their efficacy in both personal (PeAPE) and program assessment pedagogy (ProAPE) as equally high and we observed they were moderately positively correlated. MTEs feel least equipped to assist student teachers assess K-12 music learning and this was observed, along with state assessment regulation knowledge, to be a separate construct. MTEs reported feeling that their knowledge of state regulations of assessment was high. MTEs that graduated after 2008 have significantly lower levels of both personal and program assessment pedagogy efficacy than those who graduated before 2008, yet this set of individuals

report higher levels of general satisfaction in how their programs generally teach about assessment.

Research Question 5: To what extent, do the quantitative item results and the qualitative item results align with one another?

This small population sample of MTEs holds the belief that assessment is important and to an extent, we illustrate their reasoning behind this belief. The merging of the wider data did not however yield the alignment we had anticipated. We report that our participants have varied ideas about assessment yet despite this variety, they show generally high efficacious and competence beliefs. We have no reason or evidence to explain this paradoxical result. We would expect that if these individuals had high efficacious beliefs, they would have had more unified understandings about assessment practices and pedagogies.

It also seems that our MTE participants value assessment; yet, they have varied interpretations about what the word assessment means and how it is actioned. Perhaps they were implying a sense that assessment should be important due to what Maxwell (2005) might describe as participant reactivity to the researchers' framework of assessment as described in the survey items? Additionally, those participants that graduated more recently (after 2008) have slightly lower efficacy perceptions than those graduating before 2008 which does not make sense given the focus the education discipline has had on assessment in the age of accountability over the past decade.

Limitations

This study's relatively small response rate, while not unusual, represents a limitation therefore caution is needed when interpreting the findings. Our participants were mostly Caucasian (92%) which is a representation of the diversity challenge currently found in music education in higher education. Seventy percent of participants were career-line or tenured professors. We need to examine the results within the context of these participants and acknowledge the possibility that these are the only individuals afforded the time, with the interest and/or willingness to take time to participate in research, and more importantly, a research study about assessment. We speculate that career-line and tenured professors may have a particular schedule that allows them to take part in research whereas non-tenured or part-time faculty may not have equal amounts of spare time to participate.

Our small sample size did not support a comparison between full-time and part-time instructors however, future researchers could examine for those differences. We also acknowledge that most MTEs that declined our invitation most likely have markedly different perceptions about ProAPE and PeAPE than those reported here. We do not know why most MTEs declined our invitation to participate but perhaps there may be heightened faculty resistance to 'assessment' at large in the current age of performative (Ball, 2003) accountability and quality assurance.

General discussion

Our participants appear to have varied understandings about assessment yet hold mostly high efficacy beliefs. Higher education MTEs see assessment largely as a process with several distinctions; however, there is a reasonable level of inconsistency about the definition both in theory and in how they describe their practices. Most participants believe that assessment is important in the preparation of music teachers (a result supported by Orzolek, 2016). Nevertheless, variance in descriptions were frequently observed, supporting Brophy's (2017) global observation that the lexicon of assessment proves to be challenging. Our quantitative results indicated that MTEs are satisfied with how their programs teach assessment but these MTEs represent only a small sample of the MTE overall population.

We are unable to generalize other MTEs not represented in our study; we are unable to know how satisfied they may be with the way assessment is being taught at their institutions. It seems that the MTE participants in our study have high PeAPE and ProAPE which does not seem to align with the research indicating they have not had access to a great deal of coursework (Parkes & Rawlings, 2019). This also does not align with our finding in this study that MTE understanding of assessment is not quite comprehensive or unified. We are only reporting MTEs efficacy perceptions, rather than actual evidence of competence, yet we have reason to note that these beliefs might impact a teacher's effectiveness (Tschannen-Moran & Woolfolk Hoy, 2001) with respect to their teaching of assessment.

Differences were observed for those who graduated after 2008 ($n = 59$), meaning they had lower efficacy, both PeAPE and ProAPE, if they graduated more recently. Our methods do not allow for a causal explanation for this but the finding itself may be illustrative of current higher education trends, where there are supports in place. In higher education, institutionalized support teams such as "Offices of Assessment and Accreditation" can be found on most campuses in higher education. These offices are recognized for the ways in which they support higher education faculty to use assessment. Perhaps faculty with more experience at their institution have benefited from increased professional development opportunities and support for assessment than their less experienced counterparts.

There may be multiple reasons for those that graduated after 2008 to have lower levels of efficacy. Many teachers that have taught longer have not had focused training of assessment as part of the instructional process so may not recognize a need. It is not uncommon for those that have heightened levels of awareness about assessment (such as those that have graduated after 2008) to be less comfortable because assessment has received a stronger focus in higher education. Perhaps it is just not well-developed in music teacher educator preparation. Of course, it simply may be those more mature teachers have had more experience so they feel more confident. We suggest that more research is needed in this area. Additionally, higher efficacy may not directly reflect formal training of higher education faculty in coursework as the addition of formal assessment training is relatively new in music teacher educator preparation.

We had hoped that using a mixed-methods, convergent design, would yield an alignment; that an inconsistent knowledge about assessment might have also been corroborated with medium to lower efficacy scores. We had certainly anticipated a clearer answer. When we consider the sum of our findings, and the outcome of our planned convergence of methods, we see that while MTEs consider assessment to hold an important place in the preparation of music teachers, they have varied understandings about the term itself. Our current findings report generally high levels of efficacy; that they believe they can do enough to teach assessment practices, with individuals who graduated in the last 10 years feeling the least efficacious. Those who graduated before 2008 are likely have more experience in higher education and perhaps have received more on-the-job mentoring or peer support.

Conclusion

The presence of sub-skills in assessment, as seen in this study, suggest there are several elements to pedagogical assessment knowledge that are under-explored, both in our respondents and perhaps at large. Further studies should examine these sub-skills (broad teaching about assessment, providing authentic assessment experiences, giving multiple examples of assessments, and assisting future teachers with K-12 assessment design) in future research. We may only be at the front of research into this area. Assessment knowledge, and by extension assessment pedagogy knowledge and efficacy, in higher education faculty may be important to explore further, with a focus on the facilitation of professional development strategies.

References

- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215-228. doi:10.1080/0268093022000043065
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191-215. doi: 10.1037//0033-295x.84.2.191
- Biasutti, M, & Concina, E. (2018). The effective music teacher: The influence of personal, social, and cognitive dimensions on music teacher self-efficacy. *Musicae Scientiae*, 22 (2), 264-279. doi: 10.1177/1029864916685929
- Bonner, S. M. (2013). Technical quality of classroom assessments. In J. H. McMillan (Ed.) *Sage handbook of research on classroom assessment* (pp. 85-164). Thousand Oaks, CA: Sage Publications.
- Brophy, T. S. (April, 2017). Keynote introduction commentary at *Context Matters: The 6th International Symposium on Assessment in Music Education*, Birmingham City University, Birmingham, England, UK.
- Cochran, K. F., DeRuiter, J. A. and King, R. A. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44(4), 263 – 272. doi: 10.1177/0022487193044004004
- Cohen, J. W. (1998). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Creswell, J. W., & Plano-Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd Ed.). Thousand Oaks, CA: Sage Publications.
- Croy, S. R. (2018). Development of a group work assessment pedagogy using constructive alignment theory. *Nurse Education Today*, 61, 49-53. doi: 10/1016.j.nedt.2017.11.006
- Desimone, L. M. & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22. doi: 10.3102/01623737026001001
- Diaz, M. E. (2010). It is complicated: Unpacking the flow of teacher education's impact on student learning. *Journal of Teacher Education*, 6(5), 441-450. doi: 10.1177/0022487110372927
- Fautley, M. (2010). *Assessment in music education*. New York: Oxford University Press.

- Hadar, L., & Brody, D. (2010). From isolation to symphonic harmony: Building a professional development community among teacher educators. *Teaching and Teacher Education*, 26, 1641-1651. doi: 10.1016/j.tate.2010.06.015
- Kane, M., & Trochim, W. M. (2007). *Concept mapping for planning and evaluation* (Vol. 50). Thousand Oaks, CA: Sage Publications.
- Klassen, R. M., Chong, W. H., Huan, V. S., Wong, I., Kates, A., & Hannok, W. (2008). Motivation beliefs of secondary school teachers in Canada and Singapore: A mixed methods study. *Teaching and Teacher Education*, 24, 1919-1934. doi: 10/1016.j.tate.2008.01.005
- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, CA: Sage.
- Miles, M. B., & Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd Ed.). Thousand Oaks, CA: Sage Publications.
- Moè, A., Pazzaglia, F., & Ronconi, L. (2010). When being able is not enough. The combined value of positive affect and self-efficacy for job satisfaction in teaching. *Teaching and Teacher Education*, 26, 1145-1153. doi: 10.1016.j.tate.2010.02.010
- 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: GIA.
- Pallant, J. (2013). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS*. Birmingham, England: Open University Press.
- Parkes, K. A (2010). Measuring sense of efficacy in higher-education applied-music-performance teachers. *Research Perspectives in Music Education*, 14, 21-32.
- Parkes, K. A. & Rawlings, J. R. (2020). Music teacher educators' descriptions of pedagogical assessment knowledge (PAK) in in T.S. Brophy & M Haning (Eds)., *Advancing music education through assessment: Honoring culture, diversity, and practice – Selected papers from the 7th International Symposium on Assessment in Music Education* (pp. 337-348). Chicago: 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.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th Ed.). Thousand Oaks, CA: Sage Publications.

- Ruys, I., Van Keer, H., & Aelterman, A. (2011). Student teachers' skills in the implementation of collaborative learning: A multilevel approach. *Teaching and Teacher Education*, 27, 1090-1100. doi: 10.1016/j.tate.2011.05.005
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783-805. doi: 10.1016/s0742-51x(01)00036-1
- Wagoner, C. (2015). Measuring music teacher efficacy: self-identity and commitment among music teachers. *Bulletin of the Council for Research in Music Education*, 205, 27-49. doi: 10.5406/bulcouresmusedu.205.0027
- West, J. J. & Frey-Clark, M. L. (2018). Traditional versus alternative pathways to certification: assessing differences in music teacher self-efficacy. *Journal of Research in Music Teacher Education*, Online first, doi:10.1177/1057083718788035

APPENDIX A*Assessment knowledge and importance items*

Q 1. Please describe what the word “assessment” in general, means to you. For example, how do you understand the process of assessment to work (i.e., in grading, in assessing music performances, assessing student learning, assessing teacher quality, or general assessment processes that you know about).

Q 2. Please define and/or describe how you view assessment within the context of music teacher education and music teacher preparation. For example, why/why not include assessment in music teacher education?

Q3. How important is assessment to you, in the preparation of music educators?

- Extremely important (6)
- Moderately important (5)
- Slightly important (4)
- Slightly unimportant (3)
- Mostly unimportant (2)
- Extremely unimportant (1)

NB- Questions 2 and 3 are parallel.