

Exploring Rhythm-Labeling Ability and Sight-Singing Self-Efficacy Among Middle School Choral Students

Carisa Elam, Dorothy Musselwhite-Thompson,
& Brian C. Wesolowski¹

Research Perspectives in Music Education
20(1), 2-20
© 2019 The Florida Music
Education Association

Abstract

The purpose of this study was to examine rhythm-labeling ability and sight-singing self-efficacy in middle school choral students. Participants were 118 chorus students in grades six ($n = 46$), seven ($n = 37$), and eight ($n = 35$). Students completed a rhythm-labeling ability assessment and a scale to measure sight-singing self-efficacy. Assessment and data were converted to linear measures using the Multifaceted Rasch Rating Scale Model and analyzed using linear regression to see if rhythm-labeling ability predicted sight-singing self-efficacy. Results indicated that rhythm-labeling ability explained 12.6% of the variance in sight-singing self-efficacy ($F(1, 116) = 16.707, p < .001; f^2 = .14$). Implications for classroom teaching and learning are discussed.

Keywords

Assessment, Rhythm-Labeling, Self-Efficacy, Sight-Reading, Sight-Singing

Introduction

Sight-singing is a complex and multifaceted cognitive and physical process (Lehmann & McArthur 2002; Reifinger, 2012). It is a difficult musical skill due to the abstract nature of audiation and vocal pitch production (Demorest, 1998; Fine, Berry & Rosner, 2006). Sight-singing requires an ability to think in abstract terms and can be particularly challenging for adolescent students (Cutietta, 1979; Kuehne, 2007). Despite the difficulties, many choral directors understand the benefits of sight-singing, such as increased music literacy across the ensemble, more students fully participating in rehearsal, and greater independence as musicians (Floyd & Bradley, 2006; Kuehne, 2007). Researchers surveyed directors of successful sight-singing choral programs in Kentucky (Floyd & Bradley, 2006) and Florida (Kuehne, 2007) and found that a large majority of directors include sight-singing as a consistent element of their choral rehearsals both before and after preparing for formal performance evaluations. However, directors often express difficulty in consistently assessing and tracking their students' sight-singing skills (Kuehne, 2007). This study examined the psychometric qualities of new measurement tools that may allow directors to more easily collect formative data on cognitive ability associated with sight-singing achievement (e.g., rhythm labeling) and on sight-singing self-efficacy.

Researchers have noted that certain factors may predict success in sight-singing, including students' past musical experience (Demorest & May, 1995; Kopiez & In Lee, 2008), the availability of a practice period (Killian & Henry, 2005), and an inner hearing ability (Fine, Berry & Rosner, 2006; Norris, 2003; Rogers, 2013; Salisbury & Smith, 1929). Additionally,

¹ The University of Georgia

rhythmic literacy was found to predict pitch accuracy in high-school sight-singers (Henry, 2011). While there are many “best practice” opinions about the compared effectiveness between systems such as moveable *do*, fixed *do*, number systems, or neutral syllables, research has not provided any evidence of the consistent success of one method over another (Killian & Henry, 2005). Students have the potential to succeed in sight-singing through the use of any of these systems. In particular, Killian and Henry (2005) noted that the strategies used by individual singers when approaching a new piece of music may bear greater responsibility in predicting overall sight-singing achievement than the particular methodology used. To achieve sight-singing success, young singers need a systematic procedure, or a bank of practical, concrete strategies and procedures to use when they first approach a new piece of music (McPherson, 1994).

Past research has, however, examined the separate effects of pitch-reading and rhythm-reading on sight-singing ability. Brain lesion and neuroimaging mapping studies have suggested that pitch and rhythm are processed by separate neural mechanisms. While pitch processing seems to occur mostly in the right auditory cortex, rhythm processing occurs bilaterally and involves a more widespread neural network, suggesting that the brain perceives and responds to these dimensions in different ways (Peretz & Zatorre, 2005). In particular, Mishra (2016) examined the differences between rhythm sight-reading and melodic sight-singing through two meta-analyses. The meta-analysis of rhythm studies revealed a significant effect size for treatments such as rhythmic drills, systematic procedures, and bodily movements. The meta-analysis of the melodic studies revealed that melody-reading was positively affected by many variables, including creative activities, movement, solfege, and aural training. Mishra concluded that rhythmic sight-reading may be improved through the use of a counting system, rhythmic drill, and body movement, and that melodic sight-reading may be improved through collaborative practice and experience on instruments.

Concerning sight-singing pedagogy and how it relates to pitch accuracy, researchers have found positive correlations between sight-singing and a focus on tonal patterns (Reifinger, 2012), the use of hand signs or some type of pitch system (Henry, 2008; Killian, 1991; Killian & Henry, 2005), the effectiveness of melodic error detection (Killian, 1991) and targeting certain pitch skills (Henry, 2004). However, young singers often exhibit shortcomings in their abilities to perform rhythms correctly during sight-singing. According to Henry (2011), this may be due to a tendency of singers to prioritize pitch accuracy over rhythmic accuracy in a sight-singing situation. This hypothesis is based on past studies that pointed to the importance of rhythm-reading ability as a predictor of overall sight-reading ability in instrumentalists (Boyle, 1970; Elliott, 1982; McPherson, 1994) and singers (Henry, 2011).

Henry (2011) found that while pitch accuracy was overall unaffected by rhythmic tasks, a difficult pitch pattern correlated highly with rhythmic inaccuracies in high school singers. One purpose of her study was to measure pitch accuracy in the presence of difficult rhythm tasks. Overall, pitch accuracy was unaffected by rhythm tasks, meaning that students were able to perform pitches regardless of rhythmic difficulty. However, rhythmic accuracy was affected by pitch difficulty, indicating that in the presence of a difficult pitch pattern, students were likely to perform the rhythm with errors. She suggested that singers prioritize pitch over rhythm in sight-reading, and that once a singer becomes confident on the pitch of their next note, they can begin to consider the duration of that note. In other words, Henry (2011) concludes that “singers don’t [sic] care when to sing until they know what to sing” (p. 81).

Anecdotally, middle school students also seem to prioritize pitch accuracy over rhythm accuracy in sight-singing. A more developed ability to read rhythms accurately and efficiently

may contribute to a student's overall sight-singing ability. The more efficiently a student can interpret and read a rhythm, the more likely she/he may be to perform it accurately even after focusing most of their time on the pitch. This type of rhythmic literacy is supported by national chorus standards for middle school that identify musical literacy as a focus and indicate that students should be able to notate simple rhythms using a dictation method (National Coalition for Core Arts Standards, 2014). Rhythmic dictation methods such as numeric or syllabic labeling systems may help students develop an internal organization of sound and build the audiation skills that are crucial to sight-singing (Floyd & Haning, 2015).

The importance of students' beliefs in their own ability is an important factor in student achievement according to research in the field of social cognitive theory (Bandura, 1986). Students' perceptions of their own musical abilities may be connected to their ability to accomplish musical tasks (Davidson, 2010; Hartz & Bauer, 2016; McCormick & McPherson, 2003; Ritchie & Williamon, 2012). The term self-efficacy describes an individual's belief in his or her ability to perform certain tasks or behaviors. Albert Bandura (1986, 1997) noted that perceived self-efficacy can alter a person's cognitive thoughts, motivations, resiliency, and behaviors. He found that when faced with difficult tasks, people with higher self-efficacy were more likely than people with lower self-efficacy to set and commit to high goals and focus their efforts on mastery of the task. A music student's self-efficacy about a particular task or skill can affect their achievement in that area (McCormick & McPherson, 2003). From the perspective of social cognitive theory, choral students with high sight-singing self-efficacy will likely demonstrate high levels of sight-singing ability. One way to increase sight-singing self-efficacy in choral students may be to increase rhythmic literacy in an effort to combat the prioritization of pitch accuracy present during sight-singing (Henry, 2011). The purpose of this study was to examine rhythm-labeling ability and sight-singing self-efficacy in middle school choral students using a rhythm-labeling ability measure and a sight-reading self-efficacy measure. The following research questions guided this study:

1. What are the psychometric qualities (i.e., validity and reliability) of the rhythm-labeling ability and sight-reading self-efficacy measures?
2. What is the difficulty ordering of items for rhythm-labeling ability and for agreeableness in the sight-reading self-efficacy measures?
3. How much of the variance in students' sight-singing self-efficacy scores is due to their rhythm-labeling ability?

Method

Participants

Choral students from a public school in the southern United States ($N=118$; grade six, $n=46$; grade seven, $n=37$; and grade eight, $n=35$) participated in the study. The experience of the students included first-year chorus members ($n=43$), second-year chorus members ($n=43$) and third-year chorus members ($n=32$). Some students reported their participation in additional musical experiences such as band ($n=14$) and piano lessons ($n=24$). All students were enrolled in an elective chorus class with other students from their grade and received 45 minutes of chorus instruction per day under the same director. Sight-singing was a normal part of daily instruction for all students. The study took place during normally scheduled classes. All student participants completed informed consent forms as required by university and school district institutional review board protocols.

Measures

We developed a written assessment to assess rhythm-labeling ability (see Appendix A). For each item, we used publicly available 7th-10th grade All-State Chorus sight-reading selections developed for individual sight-reading evaluation. Rather than writing original examples, we chose to use All-State examples due to their strong connection to state and national standards as well as their authenticity to the students' curriculum and related classroom singing experiences. The assessment consisted of ten, two-measure examples. Students earned one point for every measure labeled correctly for a total possible score of 20. Each individual measure in the rhythm-labeling assessment was treated as its own item during the analysis process.

The 20-item sight-singing self-efficacy scale was based upon the Music Performance Self-Efficacy Scale (Zelenak, 2010). We modified the MPSES to be more representative of sight-singing self-efficacy rather than general music performance (see Appendix B). In order to address the content validity of the changes, two music education experts reviewed and approved these modifications to the items and the appropriateness of the Likert-type scale response format. Construct validity of the measure was established through a Rasch analysis to ensure that the items and response format were behaving in a manner that resulted in productive measurement (addressed below).

Procedure

Students completed the scale during class time through an online digital medium (e.g., Google Forms). The scale data were automated to a spreadsheet document via the Google Forms platform. Participants completed the sight-singing self-efficacy scale individually on the chrome books provided by the researcher to collect data about the self-efficacy of their sight-singing abilities. This sight-singing self-efficacy scale consisted of 20 statements about sight-singing self-efficacy and was shared with students through their school email accounts (see Appendix B). Students used a four-point Likert-type scale (*strongly agree*, *agree*, *disagree*, *strongly disagree*) to indicate their agreement with each statement. Two music education professionals reviewed and approved this response format.

The responses were collected and coded (*strongly agree* = 4, *agree* = 3, *disagree* = 2, *strongly disagree* = 1). Students completed the self-efficacy scale three days before the rhythm-labeling assessment to reduce any effect the rhythm-labeling assessment may have had on their perception of their sight-singing ability.

After all data from the online scale were collected, students completed the rhythm-labeling assessment. Students used pencils to label each rhythm with numerical representations using the number system with which they were familiar. For example, the rhythm quarter-quarter-eighth-eighth-quarter should have been labeled as "1, 2, 3 & 4." The instructor provided each student one copy of the assessment and allowed as much time as necessary to complete the task. On average, students completed the assessment in under 30 minutes.

Psychometric Considerations

In order to investigate the psychometric properties (i.e., validity and reliability) of the rhythm-labeling ability and sight-reading self-efficacy measures, we used the Rasch measurement model (Rasch 1960/1980). In particular, we used the Multifaceted Rasch Rating Scale Model to investigate three facets (i.e., variables): (a) students, (b) items, and (c) grade-level (Linacre, 1989). According to Engelhard (2013), the family of Rasch measurement models is an important tool for investigating the psychometric properties of psychological, behavioral, and health science measures. In particular, the benefit of the Rasch model is its requirements of

invariant measurement. According to Engelhard and Perkins (2011), the Rasch model has five specific requirements to produce meaningful measurement: (1) person-invariant calibration of test items: the calibration of the items must be independent of the particular persons used for calibration; (2) non-crossing item response functions: any person must have a better chance of success on an easy item than on a more difficult item; (3) item-invariant measurement of persons: the measurement of persons must be independent of the particular items that happen to be used for the measuring; (4) non-crossing person response functions: a more able person must always have a better chance of success on any item than a less able person; (5) unidimensionality: items must be measuring a single underlying latent variable. When the requirements of the Rasch model are adequately met, invariant measurement is achieved. When invariant measurement is achieved, the difficulty of an item does not affect the ability level of the participant, and the ability level of the participant does not affect the difficulty of the item. The requirements of invariant measurement, therefore, allow for the optimal confidence in the sample-independent measurement of latent (i.e., abstract) constructs, such as music performance achievement and self-efficacy, as well as any inferences from the measurement outcomes.

It was necessary to determine optimal confidence of sample-independent measurement for both rhythm-labeling ability and sight-singing self-efficacy in this study. Like many constructs often assessed in the field of music (e.g., performance ability, musical aptitude, affective response), rhythm labeling achievement and self-efficacy are considered to be unobservable, latent traits. In other words, they cannot be directly measured; rather, they can only be inferred based upon secondary, observable behaviors (Wesolowski, in press). The items for both measures in this study represent these secondary observable behaviors. However, the dichotomous responses (i.e., correct/incorrect) to the rhythm-labeling items and polytomous responses (i.e., strongly agree, agree, disagree strongly disagree) to the self-efficacy items are not measurements of themselves. The responses are categorical orderings of responses to the items. In order to infer measurement of rhythm-labeling ability and self-efficacy, a measurement model must be used to convert the categorical responses to linear measures. We chose the Rasch measurement model as an appropriate method for converting the ordinal level, observed responses into interval level, logit measures. The Rasch analysis was conducted using the software package *FACETS* (Linacre, 2014).

Results

Summary Statistics

Table 1 displays the summary statistics for the rhythm-labeling measure and the sight-reading self-efficacy measure. Analysis of the rhythm-labeling measure demonstrated significant differences for the student facet ($X^2_{(117)} = 922.2, p < 0.05$), item facet ($X^2_{(19)} = 653.9, p < 0.05$), and grade facet ($X^2_{(2)} = 18.7, p < 0.05$). Similarly, analysis of the sight-reading self-efficacy scale demonstrated significant differences for the student facet ($X^2_{(117)} = 1046.0, p < 0.05$), item facet ($X^2_{(19)} = 221.9, p < 0.05$), and grade facet ($X^2_{(2)} = 24.0, p < 0.05$). Significant differences across each of the three facets indicate that the measures were able to adequately separate out items based upon difficulty level and separate out students based upon achievement level (for the rhythm-labeling measure), level of agreeableness (for the sight-singing self-efficacy measure), and grade-level.

Both measures demonstrated high reliabilities of separation between students (rhythm measure, $Rel = .88$; self-efficacy scale, $Rel = .91$), items (rhythm measure, $Rel = .88$; self-efficacy measure, $Rel = .91$), and grade level (rhythm measure, $Rel = .88$; self-efficacy measure, $Rel =$

Table 1. Summary Statistics

	Rhythm-Labeling Assessment			Sight-Reading Self-Efficacy Scale		
	<u>Student</u>	<u>Item</u>	<u>Grade</u>	<u>Student</u>	<u>Item</u>	<u>Grade</u>
Measure (Logits)						
<i>M</i>	3.17	0.00	0.00	1.07	0.00	0.00
<i>SD</i>	3.39	2.71	0.47	1.40	0.76	0.20
<i>N</i>	118	20	3	118	20	3
Infit MSE						
<i>M</i>	0.95	0.93	0.94	1.01	1.01	0.99
<i>SD</i>	0.73	0.21	0.24	0.65	0.21	0.08
Std. Infit						
<i>M</i>	-0.30	-0.20	-0.70	-0.20	0.00	-0.10
<i>SD</i>	1.50	0.90	2.70	1.90	1.40	1.50
Outfit MSE						
<i>M</i>	0.97	1.09	1.15	1.05	1.11	1.09
<i>SD</i>	1.63	1.12	1.03	0.97	0.64	0.22
Std. Outfit						
<i>M</i>	0.80	0.60	0.20	-0.10	0.10	1.10
<i>SD</i>	1.20	0.70	1.10	1.70	1.70	3.00
Separation Statistics						
Reliability of Separation	0.88	0.97	0.88	0.91	0.96	0.91
Chi-Square	922.20*	653.90*	18.70*	1046.00*	221.90*	24.00*
<i>Degrees of Freedom</i>	117	19	2	117	19	2

* $p < 0.05$

.91). Interpreted similarly to Cronbach's Alpha, high reliabilities of separation for student achievement/agreeableness (i.e., the objects of measurement) indicate that the measures were able to extract differences between achievement level/agreeableness of students with high reproducibility. High reliabilities of separation for items and grade level (i.e., the agents of measurement) indicate an acceptable level of spread for the items and grades. In other words, the spread of difficulty for the items and spread of achievement for the grade levels were enough to adequately measure student achievement/student agreeableness across the resulting achievement/agreeableness levels.

Infit and outfit mean statistics (MSE) for each facet were close to the expected value of 1.00, indicating overall good fit to the model for each of the facets (Wright & Linacre, 1994). Overall, significant differences in all three facets, high reliability of separation, and good model-data fit indicate strong construct validity (item and grade facets) and predictive validity (student facets) for each measure.

Calibration of Items

Table 2 provides the calibration of items for both the rhythm-labeling measure and the sight-singing self-efficacy measure. In the rhythm-labeling assessment, each measure of music

Table 2. Calibration of Item Facet

Rhythm-Labeling Assessment							
Item Number	Observed Average	Measure	Standard Error	Infit MSE	Standardized Infit	Outfit MSE	Standardized Outfit
17	.20	4.07	.40	.64	-1.50	.63	.40
13	.20	3.91	.39	1.28	1.10	.88	.50
20	.20	3.48	.37	.80	-.80	.78	.30
16	.30	3.35	.36	.76	-1.00	.86	.30
9	.30	2.74	.33	.92	-.30	2.98	1.50
10	.30	2.74	.33	.92	-.30	2.98	1.50
15	.40	2.12	.31	1.30	2.00	1.71	1.00
14	.70	-.02	.35	1.15	.70	1.89	1.20
4	.70	-.15	.36	1.06	.30	1.19	.40
18	.70	-.15	.36	.85	-.60	.62	-.30
19	.70	-.15	.36	.73	-1.30	.58	-.40
1	.80	-1.04	.42	1.38	1.30	4.23	2.00
2	.80	-1.40	.44	.74	-.90	.44	-.10
8	.80	-1.61	.46	.82	-.50	.46	.00
3	.90	-1.82	.48	.70	-.90	.42	.00
6	.90	-1.82	.48	.97	.00	.38	.00
12	.90	-2.32	.52	1.02	.10	.32	.20
5	.90	-3.66	.65	.70	-.70	.12	.90
7	1.00	-4.14	.72	.92	.00	.16	1.40
11	1.00	-4.14	.72	.92	.00	.16	1.40

Sight-Singing Self-Efficacy Scale

Item Number	Observed Average	Measure	Standard Error	Infit MSE	Standardized Infit	Outfit MSE	Standardized Outfit
15	2.90	.32	.14	1.00	.00	.95	-.20
7	2.90	.31	.14	.89	-.80	.88	-.80
20	2.90	.30	.13	.87	-1.00	.94	-.30
12	2.90	.29	.14	.89	-.70	.83	-1.10
8	3.00	.22	.14	1.04	.30	.99	.00
19	3.00	.03	.14	1.22	1.50	1.27	1.60
16	3.10	-.03	.14	1.31	2.00	1.33	1.90
5	3.00	-.07	.14	1.03	.20	.95	-.20
9	3.10	-.07	.14	1.07	.50	1.04	.20
17	3.10	-.14	.15	1.40	2.60	1.53	2.90
18	3.10	-.16	.16	.96	-.10	1.02	.10
2	3.10	-.16	.17	1.12	.70	1.06	.40
13	3.10	-.20	.15	1.09	.60	1.05	.30
4	3.30	-.35	.14	1.07	.40	1.22	1.00
10	3.20	-.36	.15	.76	-1.60	.79	-1.20
1	3.20	-.37	.15	.72	-1.90	.63	-2.20
3	3.30	-.54	.15	.68	-2.20	.61	-2.20
14	3.30	-.55	.15	1.10	.60	1.05	.30
11	3.20	-.70	.16	.74	-1.90	.69	-1.90
6	3.50	-.78	.16	.82	-1.00	.70	-1.20

was treated as its own item. Larger logit measures on the rhythm-labeling measure represent more difficult items and smaller logit measures on the rhythm-labeling measure represent easier items. The most difficult item was Item 17 (observed average = 0.20, logit measure = 4.07). The easiest items were Item 7 and Item 11 (observed average = 1.00, logit measure = -4.14). While all items on the rhythm-labeling assessment fell within the acceptable infit range (between 0.60 and 1.40, Wright & Linacre, 1994), several notable misfit items (Outfit MSE) include items 1, 9, and 10. Misfit items indicate that the patterns of responses were too sporadic in relation to the predicted responses by the model and are therefore considered unproductive for measurement and related inferences.

Higher measures on the sight-reading self-efficacy scale represent questions on which students demonstrated the lowest amount of agreeableness (i.e., the most difficult items to agree with). Lower measure numbers represent questions on which students demonstrated the highest amount of agreeableness (i.e., the least difficult items to agree with). Students rated themselves lowest on Item 15, "Sight-reading makes me feel good" (observed average = 2.90, logit measure = 0.32). Students rated themselves highest on Item 6, "I have had positive experiences sight-reading simple music" (observed average = 3.5, logit measure = -0.78). All items except Item 17 ("I do not worry about making small mistakes during sight-reading") resulted in infit and outfit statistics within the acceptable range of 0.60 to 1.40.

Linear Regression

Table 3 shows the results of examining how rhythm labeling ability predicted sight-singing self-efficacy after misfit items 1, 9, and 10 from the rhythm-labeling task and item 17 from the self-efficacy survey were removed from calculations. The relationship was significant [$R^2 = .126$, $F(1, 116) = 16.71$, $p < .001$]. Analysis indicated a moderate positive correlation between rhythm-labeling assessment scores and reported sight-singing self-efficacy ($r = .35$) (see Figure 1). A student's predicted self-efficacy score was equivalent to $.60 + .15$ (rhythm-labeling ability) logit measures. The predicted mean sight-singing self-efficacy score for a student would increase by .15 logits for each 1.00 logit score increase on the rhythm-labeling assessment. If a student's rhythm-labeling score increased by 1.00 logit, the predicted self-efficacy score would be .15 logits higher than if the rhythm-labeling score remained the same.

Table 3. Summary of Linear Regression Analysis for Rhythm Reading Ability Predicting Sight-Singing Self-Efficacy

Variable	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>
Self-Efficacy <i>M</i> = 1.07 <i>SD</i> = 1.40	.601	.167	.000	3.608	<.001
Rhythm-Labeling <i>M</i> = 3.17 <i>SD</i> = 3.39	.147	.036	.355	4.087	<.001

Note. ($F(1, 116) = 16.707$, $p < .001$, $R^2 = .126$;
Regression equation: Self-efficacy = $.601 + .147(\text{Rhythm Assessment})$)

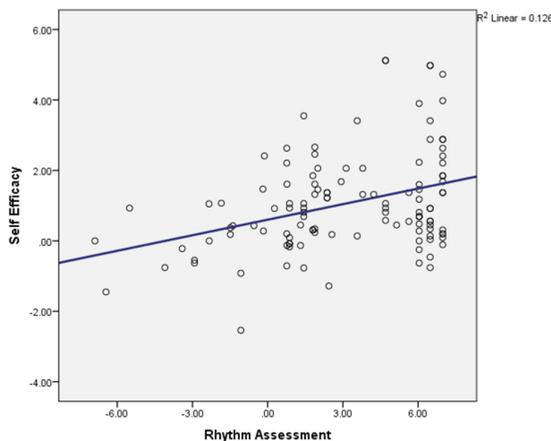


Figure 1. Scatterplot for the relationship between Self-Efficacy and Rhythm Ability; $y = .147x + .601$

Discussion

This study examined the psychometric qualities of new measurement tools and applied the Multifaceted Rasch Rating Scale Model to investigate three facets: students, items, and grade-level. The first research question addressed the psychometric qualities (i.e., validity and reliability) of the rhythm-labeling ability and sight-singing self-efficacy measure. Overall, both measures provided strong construct validity and high reliability as a result of significant differences in facets, high reliabilities of separation, and overall good model data fit for the student, item, and grade-level facets used in the measurement model. These findings provide evidence suggesting that the data generated by these tools exhibits characteristics of validity and reliability. Choral educators can use both measures to quickly and effectively collect individual data on sight-singing self-efficacy and rhythm-labeling ability. These assessments are formative in nature and can be used by educators throughout the school year to track progress. Because students can complete each assessment individually, with technology provided by the school, and without one-on-one supervision from a teacher, these measures can provide quality data without sacrificing large amounts of rehearsal time.

The second research question investigated the difficulty ordering and goodness of fit of items for each measurement tool. Student responses on each scale provided enough consistency to adequately order items by difficulty. Table 4 presents the most difficult and easiest items. The most difficult items on the rhythm-labeling assessment were measures with dotted quarter notes (Items 9, 10, 13, 16, 17, and 20). Furthermore, items that included dotted quarter notes were found to correspond exactly to the six most difficult items. Items 17 and 13, the most difficult on the assessment, were the only two items in which a dotted quarter note appears on beat two of the measure. The three items on the assessment that involve only four quarter notes (Items 5, 7, and 11) were found to be the easiest. Because the participating students were taught to read simple quarter note patterns during the first semester of sixth grade and learned dotted quarter notes on weak beats later in the school year, these findings make sense based on alignment to the program curriculum.

Items on the sight-reading self-efficacy scale were scored on a four-point Likert-type scale (*strongly agree, agree, disagree, strongly disagree*). Item difficulty measured the extent to which students, overall, agreed or disagreed with each statement. The least agreeable item was "Sight-

Table 4. Most Difficult and Easiest Items on Rhythm-Labeling and Sight-Reading Self-Efficacy Measures

Most Difficult Items	
Rhythm Assessment	Self-Efficacy Scale
17 	15 Sight-reading makes me feel good.
13 	7 People have told me that my practice efforts have improved my sight-reading skills.
20 	20 I enjoy participating in sight-reading.
Easiest Items	
Rhythm Assessment	Self-Efficacy Scale
11 	6 I have had positive experiences sight-reading simple music.
7 	11 I have received positive feedback on sight-reading activities.
5 	14 I have had positive experiences sight-reading music in a small group (5 or 6 people).

Note. Items on self-efficacy scale were ranked on a Likert-type four-point scale. A “difficult” item denotes a statement to which students most strongly disagreed.

reading makes me feel good” (Item 15). Students most strongly agreed with the statement from Item 6, “I have had positive experiences sight-reading simple music”. One possible explanation could be that the statement in Item 15 is vague and could be interpreted in many ways by students, whereas Item 6 is a more specific and agreeable statement. Perhaps the one misfitting item (Item 17: “I do not worry about making small mistakes during sight-reading”) does not represent the dimension of self-efficacy. Rather, it may reflect a dimension of performance anxiety.

The third research question examined the relationship between rhythm-labeling ability and sight-reading self-efficacy. Statistical evidence from the linear regression analysis suggested that rhythm-labeling ability and sight-reading self-efficacy are positively related ($r = .35$). The results of this study indicated that scores on the rhythm-labeling assessment accounted for 12.6% of variation in reported sight-reading self-efficacy. There were many factors not analyzed by this study that may account for more variability in the students’ scores, such as years of experience in chorus, instrumental music experience, academic achievement, interest in music, test-taking ability, etc. Future research may include a multivariate analysis of the effects of each of these factors on a student’s sight-singing self-efficacy from a more holistic perspective.

Conclusion

Limitations of this study are inherent in relying on data collected by one teacher in one music classroom. All participants received instruction under the same instructor, there was no control group, and this was the first time each of the measurement instruments were implemented in their current form. Some of these limitations could be remedied via a laboratory setting, the availability of other trained music professionals on sight, access to different students, classrooms, and school contexts, alterations to instruction, and alterations to items on the measurement instruments in order to provide better precision of student achievement and student agreeableness.

The advantages of this study's approach include the naturalized setting, the comfort of participants in the testing environment, and the insight the principal investigator had into interpretation of results due to a knowledge of the participants' general ability levels. For example, an outside examiner may be suspicious of an eighth-grade student scoring drastically lower than the sixth-grade average. But the teacher-researcher benefits from the insight that the eighth grader had just transferred into chorus class that week and can more accurately interpret the misfitting score.

It is interesting to consider the possibilities these findings may have on building sight-singing confidence in middle school choral students. The ability to sight-sing requires an efficiently executed combination of different neural networks to perceive, process, and perform pitch and rhythm simultaneously. Unlike instrumental sight-reading, singers have no tangible instrument to aid in the production of an accurate pitch. Instead, they must rely completely on their aural skills and inner-hearing ability to produce the correct pitches. While these skills are still developing, young singers may be less confident in their ability to sing the next note and may focus most of their attention on producing the correct pitch, often at the expense of maintaining rhythmic accuracy. This lack of confidence may be even more pronounced in adolescent singers, particularly when sight-singing in front of their peers.

Teaching and assessing a complex skill like sight-singing to middle school children can seem daunting and frustrating to choral educators. Fortunately, this research suggests that a simple way to increase a student's confidence in their own sight-singing may be to increase their ability to label rhythms. Further, the measures created and evaluated in this study provide educators with valid formative assessments that can be used to efficiently collect and track data on individual choral students without sacrificing rehearsal hours.

The results of this study support past research on the connection between rhythm-reading and sight-reading (Elliot, 1982; McPherson, 1994) and the potential effectiveness of emphasizing rhythm-reading strategies during rehearsal (Boyle, 1970; Henry, 2011; Killian & Henry, 2005). Results of the linear regression analysis revealed a positive correlation between student scores on each of the two measures and the predictive ability of rhythm-labeling on sight-reading self-efficacy. This insight provides justification for a more rigorous and consistent rhythm-labeling curriculum in the choral classroom. Maintaining a systematic focus on reading, labeling, and performing rhythms accurately at sight may provide choral students with the scaffolding they need when combining the complex processes of pitch- and rhythm-reading. The quicker students can accurately process rhythms, the more likely they may be to perform them correctly after focusing most of their attention on producing the correct pitches. Rhythmic accuracy is invaluable when sight-reading as an ensemble because incorrect rhythms are much more disruptive and more difficult to recover than wrong pitches.

With an understanding of the singer's tendency to prioritize pitch, teachers should consider implementing a systematic approach to sight-singing that requires choral students to

focus on the rhythm first. For example, a director can begin by having students tap their foot to the steady beat while clapping only the rhythm of a sight-singing exercise. Next, students can tap the beat while clapping and speaking the solfege, numbers, or pitch names in rhythm. After demonstrating a solid understanding of the rhythmic material, students can be allowed to add the pitch during individual practice time.

The benefit of using Rasch measurement analysis to determine the psychometric qualities (i.e., validity and reliability) of the two assessments is the ability to individually measure the fit of students, items, and grade level to the model. Both measurement tools produced high reliabilities of separation, so each facet can confidently be examined independently of the others. This allows the researcher to target any student, item, or grade level misfits for investigation. Misfit data (e.g., infit or outfit MSE scores outside of the expected 0.60 – 1.40 range (Wright & Linacre, 1994)) represent item stems that do not fit the model. In other words, the item produces data that do not cooperate with the constructed measurement trend and require further investigation. The benefit of the calibration table is the ability to quickly identify these test items, students, and grade levels that misfit the model.

Rasch analysis of the assessments used in this study indicates the need for further refinement to some of the items. Items 1, 9, and 10 on the rhythm-labeling assessment were found to be misfitting. Before future use, these items should be either removed or rewritten and retested to ensure better fit to the model. Items 9 and 10 are identical measures rhythmically (dotted-quarter note, eighth note, and two quarter notes). It is possible this redundancy may have contributed to the misfit. Further investigation into Rasch fit analysis is necessary for the researcher to hypothesize a cause for item 1 misfit.

Overall, results from the analysis indicated that choral students scored higher than expected on the rhythm-labeling assessment which suggests the presence of too many easy items. Although the measure was still found to have the ability to reliably separate scores, it may be helpful for future research to alter the assessment by removing some of the easier items (i.e., 11, 7, 5, or 12) and replacing them with more difficult rhythm patterns such as more dotted quarter note patterns on weak beats, eighth-quarter note syncopations, or dotted half notes. This edit could strengthen the targeting of the measure, specifically its ability to separate scores, allowing the researcher to more specifically differentiate between high-, middle-, and low-performing choral students.

The sight-singing self-efficacy scale proved to be a reliable measure of student perception of their own sight-singing ability. According to the Rasch analysis, each item except for number 17 was a good fit to the model and could produce a valid reflection of the latent variable of student self-efficacy. The misfit item 17 reads, “I do not worry about making small mistakes during sight-reading”. Upon further investigation, the researcher noticed that it is the only item using a negative sentence (“I do not...”). It is possible that this distinction could have contributed to the item’s misfit if students overlooked the negative and answered as if it were written “I worry about making small mistakes...” Another possibility is that this item could be unintentionally measuring performance anxiety rather than self-efficacy. Before future use, item 17 should be either rewritten or removed to provide better fit to the model. Overall, this self-efficacy scale is an efficient, quick way of checking in with middle school choral students about their sight-singing skills without having to figure out a way to evaluate them individually during very limited rehearsal time. The Google forms layout also proved to be incredibly user-friendly for the students, and the automated spreadsheet saved time for the researcher.

Collectively, the rhythm-labeling assessment and self-efficacy scale can provide teachers with a quick and effective way to glean valuable information they can use to monitor and

increase sight-reading confidence across the ensemble. In this study, choral students completed the self-efficacy scale only once, and the results represented simply a snapshot of their perception of their own sight-singing abilities at the time the scale was completed. Administering the rhythm assessment and self-efficacy scale together multiple times with the same group of choral students could perhaps be even more beneficial in tracking growth. Future research may include a pre-test/post-test study in which choral students are assessed multiple times throughout the course of one or more school years to track sight-reading self-efficacy across a period of time and examine any trends. Future research studies could also examine the use of the sight-singing scale in conjunction with other assessments, such as pitch-reading or ear training tests to empirically track possible relationships between sight-singing and other more specific skillsets.

References

- Bandura, A. (1997). Self-efficacy. *Harvard Mental Health Letter*, 13(9), 4.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Boyle, J. D. (1970). The effect of prescribed rhythmical movements on the ability to read music at sight. *Journal of Research in Music Education*, 18(4), 307–318. doi:10.1017/CBO9781107415324.004
- Cutietta, R. (1979). The effects of including systemized sight-singing drill in the middle school choral rehearsal. *Contributions to Music Education*, 7, 12–20. doi: 10.2307/24127456
- Davidson, P. (2010). The role of self-efficacy and modeling in improvisation among intermediate instrumental music students. *Journal of Band Research*, 45(2), 42–58.
- Demorest, S. M. (1998). Improving sight-singing performance in the choral ensemble: The effect of individual testing. *Journal of Research in Music Education*, 46(2), 182–192. doi: 10.2307/3345622
- Demorest, S. M., & May, W. V. (1995). Sight-singing instruction in the choral ensemble: Factors related to individual performance. *Journal of Music Teacher Education*, 43(2), 156–167. doi: 10.2307/3345676
- Elliott, C. A. (1982). The relationships among instrumental sight-reading ability and seven selected predictor variables. *Journal of Research in Music Education*, 30(1), 5–14. doi: 10.2307/3344862
- Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York: Routledge.
- Engelhard, G., Jr., & Perkins, A. F. (2011). Person response functions and the definition of units in the social sciences. *Measurement: Interdisciplinary Research & Perspective*, 9, 40–45.
- Fine, P., Berry, A., & Rosner, B. (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music*, 34(4), 431–447. doi: 10.1177/0305735606067152
- Floyd, E., & Bradley, K. D. (2006). Teaching strategies related to successful sight-singing in Kentucky choral ensembles. *Update: Applications of Research in Music Education*, 25(1), 70–81. doi: 10.1177/87551233060250010108
- Floyd, E., & Haning, M. (2015). Sight-singing pedagogy. *Journal of Music Teacher Education*, 25(1), 11–22. doi: 10.1177/1057083714539767
- Hartz, B., & Bauer, W. (2016). The effect of ear playing instruction on adult amateur wind instrumentalists' musical self-efficacy: An exploratory study. *Contributions to Music Education*, 41, 31–51.

- Henry, M. L. (2008). The use of specific practice and performance strategies in sight-singing instruction. *Update - Applications of Research in Music Education*, 26(2), 11–16. doi: [10.1177/8755123308317675](https://doi.org/10.1177/8755123308317675)
- Henry, M. L. (2004). The use of targeted pitch skills for sight-singing instruction in the choral rehearsal. *Journal of Research in Music Education*, 52(3), 206–217. doi: [10.2307/3345855](https://doi.org/10.2307/3345855)
- Henry, M. L. (2011). The effect of pitch and rhythm difficulty on vocal sight-reading performance. *Journal of Research in Music Education*, 59(1), 72–84. doi: [10.1177/0022429410397199](https://doi.org/10.1177/0022429410397199)
- Killian, J. N. (1991). The relationship between sight-singing accuracy and error detection in junior high singers. *Journal of Research in Music Education*, 39(3), 216–224. doi: [10.2307/3344721](https://doi.org/10.2307/3344721)
- Killian, J. N., & Henry, M. L. (2005). A comparison of successful and unsuccessful strategies in individual sight-singing preparation and performance. *Source Journal of Research in Music Education*, 53(1), 51–65. doi: [10.1177/002242940505300105](https://doi.org/10.1177/002242940505300105)
- Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, 10(1), 41–62. doi: [10.1080/14613800701871363](https://doi.org/10.1080/14613800701871363)
- Kuehne, J. M. (2007). A survey of sight-singing instructional practices in Florida middle-school choral programs. *Journal of Research in Music Education*, 55(2), 115–128. doi: [10.1177/002242940705500203](https://doi.org/10.1177/002242940705500203)
- Lehmann, A. C., & McArthur, V. (2002). Sight-Reading. In *The Science & Psychology of Music Performance* (pp. 135–150). Oxford University Press. doi: [10.1093/9780195138108.003.0009](https://doi.org/10.1093/9780195138108.003.0009)
- Linacre, J. M. (1989). *Many facet Rasch measurement*. Chicago, IL: MESA Press.
- Linacre, J. M. (2014). *Facets*. Chicago, IL: MESA Press.
- McCormick, J., & McPherson, G. (2003). The role of self-efficacy in a musical performance examination: an exploratory structural equation analysis. *Psychology of Music*, 31(1), 37–51. doi: [10.1177/0305735603031001322](https://doi.org/10.1177/0305735603031001322)
- McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education*, 42(3), 217. doi: [10.2307/3345701](https://doi.org/10.2307/3345701)
- Mishra, J. (2016). Rhythmic and melodic sight-reading interventions: Two meta-analyses. *Psychology of Music*, 44(5), 1082–1094. doi: [10.1177/0305735615610925](https://doi.org/10.1177/0305735615610925)
- National Coalition for Core Arts Standards (2014). National Coalition for Core Arts Standards. Retrieved from <http://www.nafme.org/my-classroom/standards/>
- Norris, C. E. (2003). The relationship between sight-singing achievement and melodic dictation achievement. *Contributions to Music Education*, 30(1), 39–53. doi: [10.2307/24127026](https://doi.org/10.2307/24127026)
- Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, 56(1), 89–114.
- Reifinger, J. L. (2012). The acquisition of sight-singing skills in second-grade general music: Effects of using solfège and of relating tonal patterns to songs. *Journal of Research in Music Education*, 60(1), 26–42. doi: [10.1177/0022429411435683](https://doi.org/10.1177/0022429411435683)
- Ritchie, L., & Williamon, A. (2012). Self-efficacy as a predictor of musical performance quality. *Psychology of Aesthetics, Creativity, and the Arts*, 6(4), 334–340.
- Rogers, M. (2013). Aural dictation affects high achievement in sight-singing, performance and composition skills. *Australian Journal of Music Education*, (1), 34–52.
- Salisbury, F. S., & Smith, H. D. (1929). Prognosis of sight-singing ability of normal school students. *Journal of Applied Psychology*, 13(5), 425–439. doi: [10.1037/h0071220](https://doi.org/10.1037/h0071220)
- Wesolowski, B. C. (in press). Item response theory and music testing. In T. S. Brophy (Ed.), *The Oxford handbook of assessment, policy and practice in music education* (pp. 479–503). New York: Oxford University Press.

- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Zelenak, M. (2010). Development and validation of the Music Performance Self-Efficacy Scale. *Music Education Research International*, 4, 31–43.

Appendix A

Rhythm-Labeling Assessment

1. 

2. 

3. 

4. 

5. 

6. 

7. 

8.

Appendix B

Sight-Reading Self-Efficacy Scale

Modified from: Zelenak, M. (2010). Development and validation of the Music Performance Self-Efficacy Scale. *Music Education Research International*, 4, 31–43.

Directions: Respond to the following statements based on your current level of sight-reading ability and experience. There are no right, or wrong, answers. Indicate to what degree you either agree or disagree with the statement by writing a number between 1 and 4 in the blank.

4 = Strongly Agree

3 = Agree

2 = Disagree

1 = Strongly Disagree

Carefully consider your choice.

- _____ 1. I have had positive experiences sight-reading music in the past.
- _____ 2. My friends think I am a good sight-reader.
- _____ 3. I have had positive experiences sight-reading in class.
- _____ 4. I have improved my sight-reading skills by watching someone I know sight-sing well.
- _____ 5. I have had positive experiences sight-reading music alone.
- _____ 6. I have had positive experiences sight-reading simple music.
- _____ 7. People have told me that my practice efforts have improved my sight-reading skills.
- _____ 8. I have had positive experiences sight-reading complicated music.
- _____ 9. I have used other music students as models to improve my sight-reading skills.
- _____ 10. I have overcome sight-reading challenges through hard work and practice.
- _____ 11. I have received positive feedback on sight-reading activities.
- _____ 12. I have used a practice routine to help me prepare for sight-reading.
- _____ 13. I am learning, or have learned, to control my nervousness during sight-reading.
- _____ 14. I have had positive experiences sight-reading music in a small group (5 or 6 people).
- _____ 15. Sight-reading makes me feel good.

- _____ 16. I have watched other students with similar musical ability as me sight-read an example and then decided whether or not I could sight-read that same example.
- _____ 17. I do not worry about making small mistakes during sight-reading.
- _____ 18. I have compared my sight-reading skills with those of other students who are similar in musical ability to me.
- _____ 19. My chorus teacher has complimented me on my sight-reading.
- _____ 20. I enjoy participating in sight-reading.