Article

AMERICAN CHORAL DIRECTORS ASSOCIATION

IJRCS

International Journal of Research in Choral Singing

The Scientific Research Journal of the American Choral Directors Association

International Journal of Research in Choral Singing
(2025) Vol. 13 101-124

The Effects of Computerized Feedback on Sight-Singing Achievement

Adam G. White[1]

Abstract

The purpose of this quantitative study was to examine the effects of computerized visual feedback provided by the *SmartMusic* interface on the sight-singing achievement of choristers (n = 77) from two suburban high schools. Using a matched group design, participants were assigned one of three groups: those who viewed feedback following their initial attempt, those who viewed feedback following their follow-up attempt, and those who did not view any feedback. Over a period of five weeks, choristers engaged in weekly sight-singing assessment sessions where they sight-sang a melody, reviewed that melody for 90 seconds, then sang that melody again. Results determined that while students made significant improvements on a melody following a sight-singing attempt, those improvements were not affected by feedback condition. These findings suggest that though feedback may be an important component in the development of sight-singing skills, the computerized feedback provided in this study was no more effective than receiving no feedback at improving sight-singing achievement. Furthermore, students were unable to transfer learning from practice with a click-track and note indicator to performance without these features so teachers should design summative assessments to match the task presented during formative assessments. This technology may be best utilized to supplement sight-singing instruction but is unlikely to supplant the work of a quality teacher.

Keywords: *sight-singing, vocal sight-reading, assessment, visual feedback*

[1] Northern Kentucky University, School of the Arts, Nunn Dr., Highland Heights, KY, 41076, USA

Corresponding author: Adam G. White; email: whitea45@nku.edu

The Effects of Computerized Feedback on Sight-Singing Achievement

The ability to sing a written melody at sight without the help of an external pitch reference is an elusive task for novice and experienced musicians alike (Demorest & May, 1995; Killian & Henry, 2005; McLung, 2008; Petty & Henry, 2014). Chorus teachers value sight-singing skill development among their choristers and feel sight-singing instruction is an important component in the rehearsal process, though more directors believe in the efficacy of sight-singing instruction than actually teach it (Farenga, 2013; Myers, 2008; Potter, 2015; Von Kampen, 2003). Some choir directors believe the process of learning sight-reading skills improves their choir's overall intonation (Floyd & Bradley, 2006). Schools of music have used sight-singing as a component for program admissions and proficiency as a requirement for graduation (Hime et al., 2014), though it is unknown how many schools have had sight-singing standards and how stringent these standards may be. In addition, adjudicated sight singing has also been a part of festivals and honor ensemble auditions in the United States. Norris (2004) found that 24 states (48%) included a formal sight-singing requirement in state-level high school choral ensemble adjudications. Several studies (Brendell, 1996; Demorest, 2001; Snider, 2007) have shown that the existence of sight-singing at festivals tended to have a positive correlation with time teachers spent on sight-singing instruction.

Researchers studied the effectiveness of different strategies for teaching and learning sight-singing, including pedagogy (Benton, 2002; Boisen, 1982; Killian, 1991; Kostka, 2000), systems (Brown, 2001; Demorest & May, 1995; Henry & Demorest, 1994; McClung, 2008), and the need for individual assessment (Demorest, 1998; Nolker, 2006). Though research on the use of various methods and solmization systems is mixed, individual assessment was found to be an effective way to improve chorister sight-singing achievement.

Researchers have been interested in the efficacy of technology as an assessment tool for several decades (Lorek, 1991; Ozeas, 1991; Platte, 1981) and advances have led to increased use of computerized assessment in choral classrooms (Hawkins, 2018; Neilsen, 2013). Researchers have also investigated the *SmartMusic* proprietary technology, including smart accompaniment and assessment features, as the focus of several instrumental music studies, investigating such topics as motivation (Gurly, 2012; Perry, 2014), attitudes toward use (Owen, 2015; Walls et al., 2013), assessment (Buck, 2008; Karas, 2005), and achievement (Flanigan, 2008). Petty and Henry, (2014) found that the use of technology for sight-singing assessment was found to be as effective as traditional methods and suggested, "While it was beyond the scope of the current study, research should be conducted to determine whether the feedback provided through the software during individual practice can impact aural skill acquisition and error detection skills" (p. 27).

Feedback

According to Kulhavy and Wager (1993), feedback "designates any information that follows a response and allows a student to evaluate the adequacy of the response itself" (p. 3).

The study of feedback has its roots in the work of Thorndike (1927; 1933), who studied how simple "right" and "wrong" feedback to student responses could affect those responses in subsequent trials. The psychologist Skinner and his study of behaviorism built on the work of Thorndike. Skinner (1965) believed that environmental stimuli either reinforced behavior or acted as punishment to diminish that behavior. Behaviorism was followed by the study of cognitivism (Gagné et al., 1981; Kulhavy & Wagner, 1993) and constructivism (Jonassen, 1990; Karagiorgi & Symeou, 2005). Each of these learning theories offered different perspectives but have all addressed the influence of feedback on learning.

Education researchers also studied the timing of feedback on learning outcomes and retention. In many cases, the timing of feedback studied was either immediate or delayed by as much as 24 hours (Clariana, 2000; Nakata, 2015). In many cases, delayed feedback contributed significantly to memory retention over immediate feedback, though students preferred immediate over delayed feedback. When offered the option of viewing delayed feedback, only 47% of participants chose to do so (Mullet et al., 2014). Researchers have not studied the timing of feedback within the context of improving a sung melody after an initial attempt.

Within music education research, the effects of learner knowledge of results (KR) has been studied (a) within teaching sequential patterns (Price, 1992), (b) during piano performance (Coffman, 1990; Banton, 1995), (c) on elementary voice development (Rutkowski & Miller, 2003; Welch, 1985), (d) while learning foreign language diction, (Steinhauer & Grayhack, 2000), (e) on success and failure attribution (Schmidt, 1995; Vispoel, & Austin, 1993), and (f) on interval identification (Jeffries, 1967). The development of pitch-recognition software makes an investigation into the efficacy of computerized KR possible within a choral music context.

Real-time computerized visual feedback was used to study singing accuracy with varying results. Welch (1985) used real-time visual feedback and KR to assist elementary children in learning an echo singing task. Groups that reviewed feedback showed greater improvement than control groups. Wilson et al. (2008) and Leong and Cheng (2014) found participants who were presented concurrent visual feedback significantly improved their singing accuracy following a training period compared to participants who did not receive feedback. Paney and Tharp (2019) found no differences among groups in a similar study. Howard (2005) found visual feedback useful during a private voice lesson setting but cautioned against displays becoming over-complicated or ambiguous. Wilson et al. (2008) suggested using a hybrid mode of instruction where teachers supplemented technology-based visual feedback with traditional methods. None of these studies used visual feedback within a sight-singing context.

A meta-analysis investigated 1,609 studies on the effects of feedback within a technology-based learning environment (Van der Kleij et al., 2015). Of primary interest in this analysis was a comparison of different feedback types, including (a) knowledge of results (KR) (correct or incorrect response indicated), (b) knowledge of correct response (KCR) (correct response indicated), and (c) elaborated feedback (EF) (explanation provided). The effect sizes of KR and KCR feedback varied based on the complexity of the learning task, with the higher-level outcomes having less effect. The effects of EF were found to be much more substantial but the variety of forms of EF varied among studies. Despite feedback being considered an essential tool in education, results on the efficacy of feedback have been mixed. When used solely to reinforce a correct response, feedback has not been found to affect achievement. Additionally, immediate feedback was effective with simple tasks but less effective in complex learning tasks (Kulhavy & Wagner, 1993). At the time of this study, no literature was found examining either the role of feedback as either KR or KCR in sight-singing achievement or the accuracy of the sight-singing feedback provided by computerized technology.

The purpose of this study was to investigate the effects of feedback on sight-singing achievement, both within a sight-singing assessment session and following a series of five sessions. I also sought to compare the accuracy of the feedback available through the *SmartMusic* assessment feature when compared to that of an expert human rater. The following questions guided this inquiry:

1. Does the presence or timing of feedback provided by the *SmartMusic* interface affect student abilities to correct errors following a sight-singing attempt?

2. Does the presence or timing of feedback provided by the *SmartMusic* interface affect student sight-singing achievement following a five-week treatment period?

3. What is the reliability of the feedback provided by the *SmartMusic* interface when compared to human expert ratings?

Method

To isolate the effects of computerized visual feedback on sight-singing achievement, this study utilized a matched-group, repeated-measure design to analyze within-session improvement and a matched pretest, posttest design to compare differences in sight-singing abilities following the treatment period. I manipulated the presence and order of the feedback provided by SmartMusic. Within each session of the treatment period (weeks 3-7), all participants attempted identical melodies twice. The within-session feedback group received feedback indicating correct and incorrect responses following the first attempt while the post-session feedback group received visual feedback following the second attempt. The control group received no visual feedback from the *SmartMusic* interface. Melody singing attempts were recorded twelve times from each participant over a period of nine weeks. See Figure 1 on the next page for a model of the research design.

The intervention in this study was the visual feedback provided by the *SmartMusic* Classic computer application loaded on an iPad Pro (10.5-inch), iOS version 12.2 (16E227) with the

Figure 1
Design. Matched group repeated measures with control

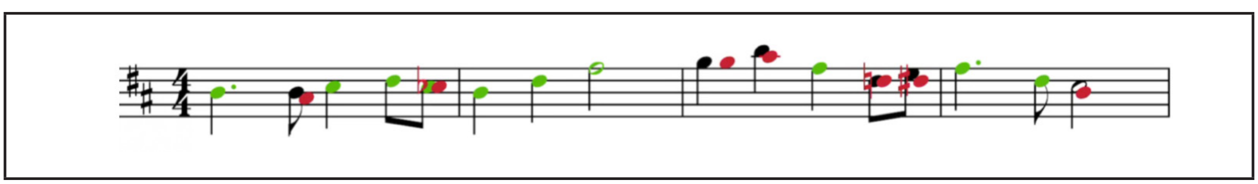
Group	Week 1	Week 3	Week 4	Week 5	Week 6	Week 7	Week 9
Within-session feedback	O_1	O_2 X O_3	O_4 X O_5	O_6 X O_7	O_8 X O_9	O_{10} X O_{11}	O_{12}
Post-session Feedback	O_1	O_2 O_3 X	O_4 O_5 X	O_6 O_7 X	O_8 O_9 X	O_{10} O_{11} X	O_{12}
Control	O_1	O_2 O_3	O_4 O_5	O_6 O_7	O_8 O_9	O_{10} O_{11}	O_{12}

O - Melody attempt

X – Feedback provided by the *SmartMusic* interface.

sight-singing instructional text, 90 Days to Sight Reading Success: A Singer's Resource for Competitive Sight-singing by McGill and Stevens (2003). All melodies were eight bars in length in 4/4 time and included the notes of the following durations: eighth, quarter, dotted quarter, and half (see Appendix A). All exercises began and ended on tonic and were in the following keys: G major, E-flat major, F major, and D major. Following a sight-singing attempt, participants in the within-session feedback (WSF) and post-session feedback (PSF) groups received visual feedback from the *SmartMusic* assessment feature (see Figure 2) that used a proprietary voice pitch-tracking algorithm. Participants in the no-feedback/control group (NFC) sang identical excerpts with the iPad microphone turned off, negating the *SmartMusic* feedback feature. Except for the presence of feedback, the *SmartMusic* interface looked identical for all participants. I used a GoPro HERO Session equipped with a 64GB ScanDisk Micro SD card as an audio and screen capture device for data analysis. An additional audio capture device, Zoom H4n, Handy Recorder was attempted for redundancy but was discontinued due to technical issues. No video recording made that included a participant's likeness, and audio recordings did not include participant names.

Figure 2
SmartMusic interface indicating correct pitches in green and incorrect in red



I established approval from the Institutional Review Board (IRB) at Northwestern University and obtained permission to conduct this study from participating schools at both the district and building levels. I made modifications to parental consent and student assent forms as needed until all parties granted approval.

Participants

Participants (*n* = 77) for this study were a convenience sample of choral students I recruited from two high schools in my professional network. Both sites were suburban high schools (grades 9-12) from a midwestern public school district. Enrollment was around 1,550 students for School A and 1,650 students for School B. Twenty-five percent of the 308 choral students recruited completed the study. Both schools had identical choral music course offerings and both used a modified block schedule. Each school had non-auditioned choirs including a tenor/bass ensemble and a treble choir. The remaining choirs were all selected by individual audition with the choir teacher. Auditioned ensembles included a select treble choir, a large mixed choir, and a small select choir, listed in order of increasing selectivity.

Procedures

Pretest/Posttest

Participants sang one of two randomly assigned melodies, A or B, during the pretest and the opposite melody during the posttest. Each test was administered by playing a screen-capture video of the SmartMusic interface that included a tonic triad (*d m s m d s, d*), thirty seconds of participant self-guided practice, another tonic triad (*d m s m d s, d*), a four-beat count off, and 50 seconds to complete the melody. The entire pretest stimulus ran for approximately one minute and 50 seconds. A click-track, quarter note indicator, and visual feedback, common features of the *SmartMusic* interface, were disabled during the pretest and posttest sessions.

Weekly Sight-Singing Sessions

A series of five, once-weekly assessment sessions began on the third week of the study. As students entered the assessment room, I verbally reviewed the assessment procedures that were as follows: (a) when I exit the room, press the microphone icon on the *SmartMusic* interface, (b) this will begin a 30 second practice period that will be preceded and followed by the tonic triad (*d m s m d s, d*), (c) sing the melody while keeping up with the click-track and quarter note indicator, (d) after completing the melody, take 60 seconds to review the melody and try to correct any errors, (e) I will re-enter the room and reset the apparatus for a second attempt, (f) when I exit the room press the microphone icon on the *SmartMusic* interface, (g) sing the melody a second time and try to improve upon your initial attempt, and (h) exit the room. When I entered the room following the first attempt and 60 seconds of practice, I enabled the microphone feature for participants in the post-session feedback group and disabled it for those in the within-session feedback group, and left it disabled for those in the NFC group. Each session took approximately four minutes and 35 seconds.

Scoring

I scored all pretest and posttest melody attempts (*n* = 154) using the following procedures:

Each eight-measure sight-singing sample was divided into two, two-count chunks (counts 1-2 and counts 3-4) for a total of 16 chunks. Each chunk was then awarded one point for the correct notes and one point for the correct rhythm for a total of 32 possible points per sample. If any error was made within a chunk, the entire chunk was awarded a zero. All scores for both pitch and melody were converted into a proportion of correct chunks per attempt. A random sample of approximately 20% (n = 30) pretest and posttest melodies were scored by an additional expert rater to establish reliability. A proportion of agreements divided by agreements plus disagreements (Madsen & Madsen, 1970), yielded a proportion of agreement of (.925) for pitch and (.856) for rhythm.

I also scored all weekly assessment session attempts (n = 770). In order to account for the unique nature of the *SmartMusic* interface, I used a different scoring method than I used for the pretest. I awarded a single point for each correct pitch and a point for each correct rhythm, similar to other studies (Henry 2004; 2011). Unique to this study, however, participants were required to stay within a quarter step, sharp or flat, of the written pitch and rhythms to be aligned with the click track and quarter note indicator to be marked correct. See Appendix A for a complete list of scoring guidelines. In order to ascertain reliability, an additional expert rater scored 30% (n = 235) of the sight-singing attempts, selected at random. Using a formula of agreements divided by agreements plus disagreements, I was able to determine a proportion of agreement for pitch (.908) and rhythm (.852) for the melodies in weeks one through five.

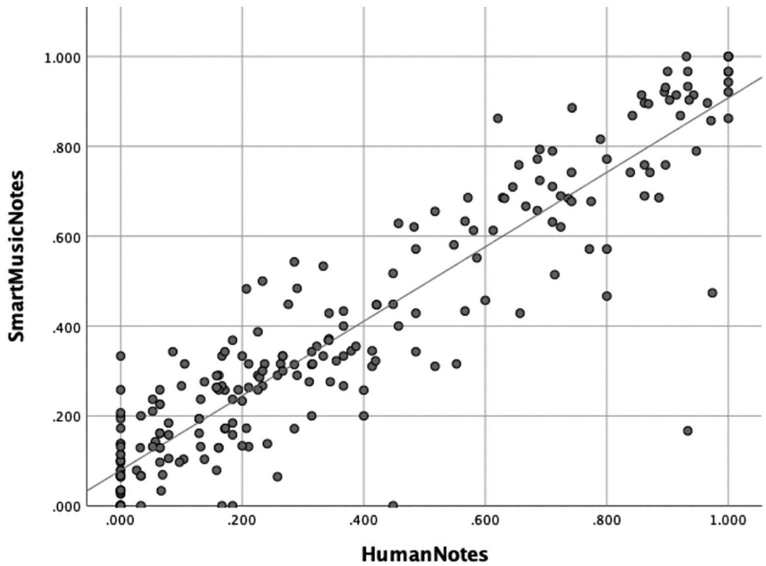
In order to determine the reliability of my pitch and rhythm proportion scoring, I ran a Type A (absolute agreement) intraclass correlation coefficient (ICC), comparing my scores to the additional rater's scores, revealing a high degree of reliability. The single measures ICC for pitch scores was .939 with a 95% confidence interval from .923 to .953, $F(244, 244) = 32.337$, $p < .001$. The single measures ICC for rhythm was .904 with a 95% confidence interval from .569 to .959, $F(244, 244) = 33.349$, $p < .001$. As the ICC for both pitch and rhythm fell within the "*excellent reliability*" range (Koo & Li, 2016), I proceeded to use my full set of scores without modification.

SmartMusic Reliability

In order to address the reliability of the SmartMusic interface compared to human raters, I compared the visual feedback captured by the video apparatus to my scores. Because the apparatus did not account for the possibility of separate pitch and rhythm scores, I considered a note correct when I scored both the pitch and rhythm as correct. I considered the note correct from the *SmartMusic* interface when the notehead was green instead of black or red. I ignored all other extraneous marks on the feedback. See Figure 3 on the next page for a comparison of human and computerized scoring.

Figure 3

Scatterplot comparison of SmartMusic and human raters scoring proportions



Results

Question 1: Does the presence or timing of feedback provided by the *SmartMusic* interface affect student abilities to correct errors following a sight-singing attempt?

I endeavored to discover if the presence or timing of feedback affected participants' ability to improve accuracy on the performance of a melody following an initial sight-singing attempt by developing the following three-level panel data regression model, regressing several independent variables on the follow-up attempt composite score. The first level model included follow-up attempt composite scores as the outcome variable and the initial attempt composite scores as a predictor variable:

$$FollowupAttemptComp_i = \alpha + \beta_1 InitialAttemptComp_i + \varepsilon$$

The second model added dummy variables for the within-session feedback group and the post-session feedback group:

$$FollowupAttemptComp_i = \alpha + \beta_1 InitialAttemptComp_i + \beta_2 WSF_i + \beta_3 PSF_i + \varepsilon$$

The final model added other dichotomous predictor variables, including choir selection, school, and voice range:

$$FollowupAttemptComp_i = \alpha + \beta_1 InitialAttemptComp_i + \beta_2 WSF_i + \beta_3 PSF_i + \beta_4 STC_i + \beta_5 LMC_i + \beta_6 SMC_i + \beta_7 School_i + \beta_8 Voice_i + \varepsilon$$

In all three models, the initial attempt was the strongest significant predictor of success on the follow-up attempt. Other factors were also significant, including choir selection and school attendance. Research condition and voice range were not significant predictors of the outcome. See Table 1 for the regression analysis.

Table 1
Summary of Hierarchical Regression Analysis for Variables Predicting Follow-up Melody Accuracy (n = 384)

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	β	*B*	*SE B*	β	*B*	*SE B*	β
Constant	.121	.014		.132	.017		.132	.023	
Initial attempt	.891	.026	.871**	.895	.026	.874**	.815	.031	.796**
WSF[a]				-.030	.019	-.044	-.028	.019	-.042
PSF[b]				-.012	.019	-.018	-.010	.019	-.014
Select Treble Choir[c]							.027	.027	.084*
Large Mixed Choir[d]							.080	.021	.108**
Small Mixed Choir[e]							.088	.024	.120**
School[f]							-.041	.016	-.064*
Treble Voice[g]							.006	.017	.010
R^2		.757			0.760			.777	
F for change in R^2					1.411			5.97*	

Note:
[a]Within-session feedback condition = 1, Post-session feedback and control = 0.
[b]Post-session feedback condition = 1, Within-session feedback and control = 0.
[a, b]Control group is constant = 0
[c, d, e]Non-auditioned choir is constant = 0
[c]Treble Choir = 1, dLarge Mixed Choir = 1, eSmall Mixed Choir = 1
[f]School A = 1, School B = 0
[g]Treble Voice = 1, Tenor/Bass Voice = 0
$^{*}p < .05.$ $^{**}p < .001.$

I created the following two-level regression model to determine any effect of variables on possible gains between attempts. See Table 2 for the regression results.

Table 2
Summary of Hierarchical Regression Analysis for Variables Predicting Gains in Composite Scores Between Initial and Follow-up Attempts (n = 384)

	Model 1			Model 2		
Variable	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Constant	.087	.013		.083	.022	
WSF[a]	-.038	.019	-.115*	-.040	.020	-.118*
PSF[b]	-.005	.020	-.015	-.007	.020	-.019
Select Treble Choir[c]				.030	.027	.063
Large Mixed Choir[d]				.042	.021	.0112*
Small Mixed Choir[e]				.010	.021	.027
School[f]				-.031	.016	-.099
Treble Voice[g]				.007	.018	.021
R^2		.012			.035	
F for change in R^2					0.133	

Note:
[a]Within-session feedback condition = 1, Post-session feedback and control = 0.
[b]Post-session feedback condition = 1, Within-session feedback and control = 0.
[c, d, e]Non-auditioned choir is constant = 0
[c]Treble Choir = 1, [d]Large Mixed Choir = 1, [e]Small Mixed Choir = 1
[f]School A = 1, School B = 0
[g]Treble Voice = 1, Tenor/Bass Voice = 0
$^{*}p < .05$. $^{**}p < .001$.

The first model included composite gain scores as the outcome variable with dummy variables representing within-session feedback and post-session feedback groups:

$$GainScoreComp_i = \beta_1 WSF_i + \beta_2 PSF_i + \varepsilon$$

The second model added the remaining dichotomous predictor variables, including choir selection, school, and voice range:

$$GainScoreComp_i = \beta_1 WSF_i + \beta_2 PSF_i + \beta_3 STC_i + \beta_4 LMC_i + \beta_5 SMC_i + \beta_6 School_i + \beta 7 Voice_i + \varepsilon$$

I applied a Pearson's *r* correlation between mean initial composite scores and mean composite gains to determine if a relationship existed between how well participants scored on their first attempt and how much they improved during their second attempt. A weak, non-significant positive relationship was found $r = .130$, $p = .258$. When I applied a cubic line-of-fit to a scatterplot, Figure 4, comparing average initial attempts and average gains among all participants, an inverted-U shaped line was revealed ($R^2 = .255$) that better accounted for variance in the data than a linear line ($R^2 = .017$).

To determine if differences in gains changed by group over time, I compared scores from each week using a two-way repeated measures ANOVA. During week one, participants scored significantly higher on the follow-up (M = .416, SD = .035) attempt than they did during the initial (M = 293, SD = .030) attempt $F(1.00, 75.000) = 51.618$, $p < .001$, $\eta_p^2 = .408$ using Greenhouse-Geisser corrected degrees of freedom. Differences among groups were non-significant and showed nearly parallel improvement. Week 2 showed significant gains for each group, though less pronounced than Week 1, with initial attempt scores significantly higher than the Week 1 initial scores. As with Week 1, pitch and rhythm scores were significantly higher during the follow-up attempt $F(1.00, 72.000) = 26.477$, $p < .000$, $\eta_p^2 = .269$. Pairwise post-hoc analysis found no significant differences between initial and follow-up attempts by condition. A two-way repeated measures ANOVA was used to compare Week 3 initial pitch and rhythm scores to follow-up attempts. Unlike weeks 1, 2, 4, and 5, I found no main effect between the initial and follow-up attempt $F(1.000, 75.000) = 2.861$, $p = .095$, $\eta_p^2 = .037$. The ANOVA analysis from Week 4 determined that follow-up scores were significantly higher than those during the initial attempt $F(1.000, 74.000) = 16.665$, $p < .001$, $\eta_p^2 = .184$. A post-hoc pairwise comparison found no differences among groups, however. The Week 5 ANOVA revealed significant differences between initial pitch and rhythm scores and follow-up scores, $F(1.000, 74.000) = 12.389$, $p = .001$, $\eta_p^2 = .143$, with no significant differences by condition.

Question 2: Does the presence or timing of feedback provided by the *SmartMusic* interface affect student sight-singing achievement following a five-week treatment period?

To compare possible student growth in sight-singing scores among groups, I performed a two-by-two repeated measures ANOVA of pretest and posttest pitch and rhythm scores with condition as a between-subjects factor. As the assumption of sphericity could not be met, a Greenhouse-Geisser correction was applied, revealing no significant differences between pretest and posttest composite scores F(1.000, 71.000) = 2.106, p = .151, η_p^2 = .029. Additionally, between-subjects comparisons revealed no significant differences between groups F(2, 71) = 2.492, p = .090, η_p^2 = .066. See Table 3 for a comparison of pretest and posttest means by group and melody. A comparison of composite gains from pre- to posttest by condition revealed positive gains in the within-session feedback group (ΔM = .046) and no-feedback group (ΔM = .094) but negative gains in the post-session feedback group (ΔM = -.036). None of these differences were statistically significant.

Table 3

Pretest/posttest composite score comparison by melody and condition

Composite score	Pretest		Posttest	
	M	*SD*	*M*	*SD*
Within-session Feedback				
Melody A	.264	.085	.385	.089
Melody B	.385	.089	.341	.089
Post-session Feedback				
Melody A	.277	.085	.216	.102
Melody B	.362	.097	.351	.089
No feedback/control				
Melody A	.214	.082	.259	.086
Melody B	.308	.082	.451	.086

Note: Composite scores are reported as proportion correct notes and rhythms

Question 3: What is the reliability of the feedback provided by the SmartMusic interface when compared to human expert ratings?

I sought to examine the reliability of the feedback provided by the SmartMusic computer application when compared with human expert scoring. Visual feedback was presented once weekly to participants in the within-session feedback and post-session feedback groups. Approximately one-third (n = 237) of all attempts received feedback. Each note was considered accurate on the visual feedback when the notehead was colored green. I disregarded all other markings provided by the feedback. Notes were considered accurate by the human rater when both the pitch and rhythm were judged to be correct.

I analyzed reliability using a proportion of agreements divided by agreements plus disagreements (C. K. Madsen & C. H. Madsen, 1970). The proportion of agreement between *SmartMusic* and my scores (n = 237) had a mean of 0.841 (SD = .124). Scores ranged between full agreement (1.0) and low agreement (0.167). The 95% confidence interval was between 0.825 and 0.857. Figure 3 provides a scatterplot of this relationship.

Summary

High school chorister volunteers (n = 77) from two suburban public schools completed a five-week sight-singing assessment session that was preceded by a pretest and followed by a posttest. A comparison of pretest and posttest scores revealed a slight, non-significant improvement among participants from pretest to posttest but found no significant differences by condition. Analysis of sight-singing scores on initial attempts during each weekly session revealed significant improvement between weeks one and three and weeks two, four, and five. A comparison of each initial attempt to follow-up attempts revealed significant gains for weeks one, two, four, and five. Participant gains during week three were not significant. A non-significant, negative relationship was found among singing scores with participants in the WSF group and the PSF group when compared to those in the NFC group. Voice part was not a significant predictor of accuracy, but school and choir placement predicted higher achievement. A comparison of *SmartMusic* feedback and human scoring revealed a very strong positive correlation when comparing the proportion of correct notes $r(235) = .923$, $p < .001$. Analysis of agreements divided by possible agreements between human and *SmartMusic scoring* revealed 84.1% consistency of scores.

Discussion

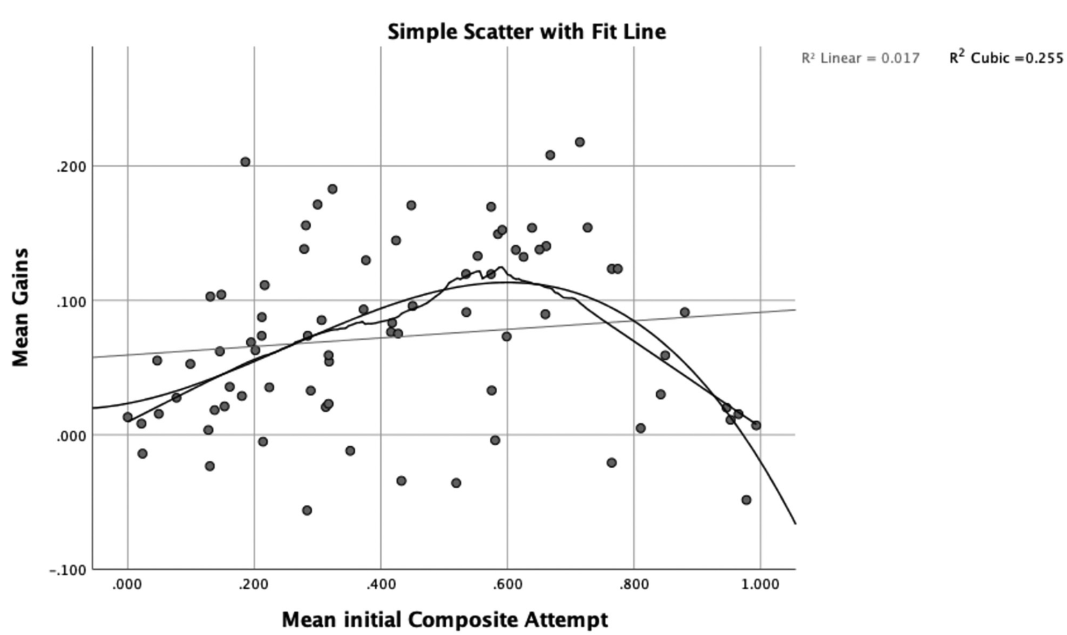
With question one, I sought to ascertain if the presence or timing of feedback affected participants' ability to improve accuracy on the performance of a melody following an initial sight-singing attempt. The three-model regression analysis listed in Table 1 compared the scores of the follow-up attempt for every melody during the five-week treatment period to a series of predictor variables. Not surprisingly, the results revealed that the greatest predictor of sight-singing achievement on the second attempt was the score of the initial attempt. This model revealed no significant differences by group assignment, suggesting feedback had no discernible effect in overall sight-singing achievement. Significant predictors were found among the choir enrollment; participants who were enrolled in more select choirs were more likely to have higher sight-singing scores on the follow-up attempt. This suggests that students placed in more select choirs were more likely to demonstrate sight-singing acumen, corroborating the findings of Demorest and May (1995).

An additional regression model (Table 2) compares participants' gains made during each session of the five-week treatment period to determine if group assignment and other variables predicted differences. The first model compared group assignment and a small but significant negative relationship ($\beta = -.115$, $p < .05$) among participants in the WSF group when compared to the control group. Participants in the PSF group did not show any significant differences in gains when compared to the control. These findings reveal that students who were not given feedback prior to a second attempt showed significantly greater gains than those who received feedback, though it should be noted that participants in the within-feedback group scored consistently higher on the initial attempt, possibly limiting their potential growth when compared to the other groups.

A comparison of gains between the initial and follow-up attempts revealed significant improvement with a large effect size for each group every week except Week 3. Differences among groups were non-significant and followed mostly parallel gains each of the five weeks of the study. The scatterplot (Figure 4) displaying the cubic relationship between average initial composite scores and average gains for all participants during weeks 1-5 further displays the need for level-appropriate melodies. Participants who averaged below 20% during their initial attempts were less likely to show as much improvement as those who scored between 20% and 80%, despite having the greatest opportunity for gains, corroborating the findings of Killian and Henry (2005). Those who averaged above 80% likely reached a ceiling effect as they approached 100% correct. Additionally, the feedback provided by *SmartMusic* was of no benefit to participants in the WSF group when compared with the other groups, even among those participants who scored below 20%.

Figure 4

Scatterplot of Average Initial Attempt Score and Average Gains



Note: Composite and gain scores are reported in proportion of correct notes and rhythms

With the second research question, I attempted to determine whether the presence or timing of computerized feedback affected sight-singing achievement following a five-week treatment period. Prior research has established the efficacy of individual assessment in improving sight-singing achievement (Demorest, 1998; Henry, 2014; Henry & Petty, 2014).

However, under these conditions, significant improvements in sight-singing achievement were not found when comparing pretest and posttest scores. Furthermore, group assignment had no significant effect on sight-singing scores. Several factors may have contributed to this lack of improvement. The pretest and posttest procedures differed from those experienced by the participants during the weekly sight-singing sessions. While the interface and initial practice time were identical, the click-track and quarter note indicator were disabled. It may be reasonable to assume that after 10 melody attempts with these features that some participants became accustomed to, if not reliant upon, these features when attempting a melody.

It remains unclear if either the frequency of sight-singing sessions, whether daily, weekly, or monthly, or duration of the treatment period, longer than five weeks, would have changed these findings. It should also be noted that though participants in the within-session feedback group scored the highest during the initial attempt each of the five weeks of the study, differences between groups maintained a parallel motion, suggesting that differences in scores were more likely the result of differences between groups that existed at assignment rather than the result of the research condition.

With research question three, I explored the reliability of *SmartMusic* feedback in comparison to manual scoring. Agreement on note accuracy was notably high, especially considering the complexities of the human voice. Alignment with an additional human rater reached 91.5%, while agreement with the *SmartMusic* system was 84.1%. These results suggest that choir teachers can reasonably view the feedback as a useful indicator of choristers' sight-singing abilities. However, caution is warranted when considering the assessment feature for formal grading, particularly in high-stakes contexts. The software demonstrated sensitivity to ambient noise from nearby rehearsal spaces and had difficulty evaluating quieter singers.

Feedback

Sight-singing is a complex task. As a result, if one of the component skills of sight-singing is missing, the singer is unlikely to be successful. As the ability to read, understand, and audiate written notation is a key sight-singing skill (Fournier et al., 2017; Vujović & Bogunović, 2012), it is logical to assume that if a student lacks a basic understanding of written notation, feedback using that notation is likely meaningless. Additionally, students who struggle to sing accurately, so called "poor-pitch singers," (Pfordresher & Brown, 2007) may understand the notation, but not sing with enough accuracy to produce a response from the visual feedback apparatus. Though not common, some participants in this study sang the correct solfege syllables in rhythm but failed to sing the pitches accurately.

Additionally, this study used a single, one-model approach as all participants sang the same melodies regardless of ability level and as a result, neglected the use of feedback to provide information informing future instruction (Fautley, 2010). Participants who struggled to maintain key, for example, received feedback that may not have presented useful information. A more effective use of this feedback may be to assign shorter melodies with a narrower range until the participant finds some success. Furthermore, participants who were able to perform the melody correctly on the first attempt could not make any improvements so the feedback, though confirming accuracy, did not provide assistance. A failure to use the feedback to alter the assessment may have narrowed its possible effectiveness to a small range of participants.

The timing of the feedback for this study, though varied by condition, was provided immediately following a melody singing attempt. Research has suggested that delayed feedback can reinforce learning and retention (Clariana, 2000; Nakata, 2015). It is possible that had the visual feedback been delayed 24 hours, it may have been more useful. Demorest (1998) utilized delayed feedback as participants in that study were given general comments and approximate scores after the attempts had been scored. It is possible that presenting students with that information after a period of time could have enhanced learning.

Additionally, this study's design did not allow for a fourth group, one where participants received feedback twice, once following the initial attempt and once after the follow-up attempt, or a fifth group where students were allowed to practice sight-singing without being recorded, thus eliminating the assessment piece. Another possibility is that the feedback was too overwhelming and failed to present a clear path to improvement. Figure 5 displays feedback from a sight-singing attempt by a student who was very close to being accurate but was either late, in the wrong key, or a combination of both. In this case, the feedback failed to provide a clear description of what went wrong or how the melody attempt could be improved.

Figure 5

Computerized feedback may not offer useful information

Limitations

Due to the quasi-experimental nature of this study and limited scope of the participants, findings of this study cannot be generalized to the population as a whole. As with many sight-singing studies that ask for volunteers (Demorest, 2001), there is a high likelihood of selection bias among these participants. Findings are also specific to the procedures detailed above and any change in those procedures would have the potential of producing different results. It is also possible that five weeks was simply too brief a timeline to achieve significant results from pre- to posttest. Findings of this study relating to feedback were specific to the visual feedback provided by the stimulus. It should not be assumed that because the feedback used in this study did not produce differences among groups that student access to feedback is not still an important part of the learning process, though the limitations of this feedback offers teachers and researchers the opportunity to keep looking for effective ways of giving students useful information that leads to musical growth.

Future research

It is possible that discernible differences would have been found among groups in this study if participants had been assigned sight-singing melodies that presented an appropriate level of challenge for their abilities. In such a situation, the feedback may have become more beneficial to a greater number of participants. Researchers have also suggested that self-efficacy and the belief that time-on-task will be productive play an important role in how research participants engage with feedback (Madsen & Duke, 1985; Timmers et al., 2013). Future researchers could design and test sight-singing methods that track student self-efficacy for sight-singing and explore different feedback models that highlight improvement.

During this study, participants were only allowed to use the *SmartMusic* interface during weekly in-class assessment sessions. It is unknown if students given free access to the software would engage with it outside of rehearsal. It is possible that students who were motivated to learn sight-singing skills would practice on their own. Future researchers might examine how students self-regulate during sight-singing practice when using technology. Additionally, researchers have yet to quantify what level of sight-singing skill allows for chorister independence and under what conditions they are indeed independent.

This study is the first among the extant research literature I reviewed that allowed participants to attempt a melody again following an initial sight-singing attempt. It was encouraging that many students, regardless of feedback condition, diagnosed errors and corrected them in a subsequent attempt. The design of this study did not provide any insight into the processes with which the students undertook, either with or without feedback, to correct mistakes. Researchers could design a study where participants talk aloud while reflecting on a sight-singing attempt or while preparing a follow-up attempt. Eye-tracking technology may also offer insight into student interactions with feedback.

The pitch recognition software used by *SmartMusic* likely demonstrated enough reliability that the potential for building an interactive platform exists. This technology might be useful for building a scaffolded interface that adjusts difficulty as participants improve using targeted pitch skills (Henry, 2004) and would be worthy of future study. The potential exists to create sight-singing software that is more appealing to choristers. Software developers and researchers could use an interactive video game model that balances challenge with the user's skill level.

Implications

Student scores found in the weekly sight-singing assessment sessions suggest that individual assessment did improve sight-singing achievement, corroborating earlier research (Demorest, 1998; Henry, 2015; Petty & Henry, 2014). This study provides evidence of the importance of adapting assessment difficulty to meet student abilities appropriately. Feedback should not be unidirectional as was the case in this study. Teachers who use technology like this should continually monitor and respond to student performance by altering instruction and future assessments. This technology may be best utilized as a supplement to sight-singing instruction, but will not replace a quality teacher.

Students need to be assessed in the same manner in which they practiced sight-singing. One possible reason for students' lack of improvement from the pretest to the posttest is that the posttest procedures did not match the assessment session procedures, or, worded differently, the summative assessment procedures did not match the formative assessment practice. Students should not be expected to sight-sing individually when their only practice was in a group setting. Furthermore, they should not perform without a metronome and quarter-note indicator if they were a regular part of their instruction. The potential exists for teachers to use technology like SmartMusic to facilitate individual assessment and curate individual sight-singing attempts electronically, making individual assessment more efficient. The need for teachers to be able to engage students in asynchronous instruction and assessment has become very pertinent (Chrysostomo & Triantafyllaki, 2020) so teachers should continue to explore the opportunities for students to engage with technology as a means of individual assessment.

As long as choirs continue to perform music written in traditional Western notation, sight-singing will be an important skill in the development of choristers' musical independence. The visual feedback used in this study emerged as a potentially viable tool to supplement chorus directors' ability to teach sight-singing. The voice-pitch recognition software used for this study has potential benefits, but it is incumbent upon software developers and teachers to use it in a manner that promotes student learning. As we continue to make individual assessment more effective, choir directors will be able to give choristers the best tools possible to make music independently and enjoy a lifetime of reading choral music.

References

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music, 23*(1), 3-16.

Benton, C. W. (2002). *A study of the effects of metacognition on sight-singing achievement and attitudes among students in a middle school choral music program* [Doctoral dissertation, Shenandoah University].

Boisen, R. (1981). The effect of melodic context on student's aural perception of rhythm. *Journal of Research in Music Education, 29*(3), 165-172.

Brendell, J. K. (1996). Time use, rehearsal activity, and student off-task behavior during the initial minutes of high school choral rehearsals. *Journal of Research in Music Education, 44*(1), 6-14.

Brown, K. D. (2001). *Effects of fixed and movable sightsinging systems on undergraduate music students' ability to perform diatonic, modulatory, chromatic, and atonal melodic passages* [Doctoral dissertation, University of Oregon].

Buck, M. W. (2008). *The efficacy of SmartMusic® assessment as a teaching and learning tool* [Doctoral dissertation, University of Southern Mississippi].

Clariana, R. B. (2000). Feedback in computer-assisted learning. *NETg University of Limerick Lecture Series.*

Coffman, D. D. (1990). Effects of mental practice, physical practice, and knowledge of results on piano performance. *Journal of Research in Music Education, 38*(3), 187-196.

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.

Demorest, S. M. (2001). *Building choral excellence: Teaching sight-singing in the choral rehearsal.* Oxford University Press.

Demorest, S. M., & May, W. V. (1995). Sight-singing instruction in the choral ensemble: factors related to individual performance. *Journal of Research in Music Education, 43*(2), 156-167.

Farenga, J. (2013). *Arizona high school choral educators' attitudes toward the teaching of group sight singing and preferences for instructional practices.* [Doctoral dissertation, Arizona State University].

Fautley, M. (2010). *Assessment in music education.* Oxford University Press.

Floyd, E. & Bradley, K. D. (2006). Teaching strategies related to successful sightsinging in Kentucky choral ensembles. *Update: Applications of Research in Music Education, 25*(1), 70-81.

Fournier, G., Moreno Sala, M. T., Dubé, F., & O'Neill, S. (2017). Cognitive strategies in sight-singing: The development of an inventory for aural skills pedagogy. *Psychology of Music*, 0305735617745149.

Gagné, R. M., Wager, W., & Rojas, A. (1981). Planning and authoring computer-assisted instruction lessons. *Educational technology, 21*(9), 17-26.

Hawkins, J. A. (2018). *Secondary choral music educators' use of technology-assisted assessment tools* [Doctoral dissertation, University of Illinois at Urbana-Champaign]. Henry, M. (2015). Vocal sight-reading assessment: Technological advances, student perceptions, and instructional implications. *Update: Applications of Research in Music Education, 33*(2), 58-64.

Henry, M. L. (2013). The effect of key on vocal sight-reading achievement. *Texas Music Education Research*, 3, 8.

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.

Henry, M. (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.

Henry, M. L., & Demorest, S. M. (1994). Individual sight-singing achievement in successful choral ensembles. *Update: Applications of Research in Music Education, 13*(1), 4-8.

Hime, L. J., Miksza, P., & Hunsucker, A. (2014). Admissions criteria and upper-divisional ratings as predictors of student teaching achievement in music education. *Bulletin of the Council for Research in Music Education*, (200), 7-22.

Howard, D. M. (2005). Technology for real-time visual feedback in singing lessons. *Research Studies in Music Education, 24*(1), 40-57.

Jeffries, T. B. (1967). The effects of order of presentation and knowledge of results on the aural recognition of melodic intervals. *Journal of Research in Music Education, 15*(3), 179-190.

Jonassen, D. H. (1990). Thinking technology: Toward a constructivist view of instructional design. *Educational Technology, 30*(9), 32-34.

Karagiorgi, Y., & Symeou, L. (2005). Translating constructivism into instructional design: Potential and limitations. *Journal of Educational Technology & Society, 8*(1), 17-27.

Karas, J. B. (2005). *The effect of aural and improvisatory instruction on fifth-grade band students' sight reading ability.* [Doctoral dissertation, The University of Nebraska-Lincoln].

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.

Killian, J. N. & Henry, M. (2005). A comparison of successful and unsuccessful strategies in individual sight singing preparation and performance. *Journal of Research in Music Education, 53*(1), 51-65.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

Kostka, M. J. (2000). The effects of error detection practice on keyboard sight-reading achievement of undergraduate music majors. *Journal of Research in Music Education, 48*(2), 114-122.

Kulhavy, R. W., & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. *Interactive Instruction and Feedback*, 320.

Leong, S., & Cheng, L. (2014). Effects of real-time visual feedback on pre-service teachers' singing. *Journal of Computer Assisted Learning, 30*(3), 285-296.

Lorek, M. J. (1991). Computer analysis of vocal input: A program that simulates college faculty sight singing evaluation. *Computers in Music Research, 3*, 121-138.

Madsen, C. K. & Duke, R. A. (1985). Perception of approval/disapproval in music. *Bulletin of the Council for Research in Music Education, 85*, 119-130.

Madsen, C. K., & Madsen, C. H. (1970). *Experimental research in music*. Prentice Hall.

McClung, A. C. (2008). Sight-singing scores of high school choristers with extensive training in movable solfège syllables and Curwen hand signs. *Journal of Research in Music Education, 56*(3), 255-266.

McGill, S., & Stevens Jr., H. M. (2003). *90 days to sight reading success: A singer's resource for competitive sightsinging.* AMC Publications.

Mullet, H. G., Butler, A. C., Verdin, B., von Borries, R., & Marsh, E. J. (2014). Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *Journal of Applied Research in Memory and Cognition, 3*, 222-229.

Myers, G. C. (2008). *Sight-singing instruction in the undergraduate choral ensembles of colleges and universities in the Southern Division of the American Choral Directors Association: Teacher preparation, pedagogical practices and assessed results* [Doctoral dissertation, The University of North Carolina at Greensboro].

Nakata, T. (2015). Effects of feedback timing on second language vocabulary learning: Does delaying feedback increase learning? *Language Teaching Research*, *19*(4), 416-434.

Nolker, D. B. (2006). The relationship between large ensemble sight-reading rating and the individuals' sight-singing success. *Missouri Journal of Research in Music Education, 43*, 3-16.

Norris, C. E. (2004). A nationwide overview of sight-singing requirements of large group choral festivals. *Journal of Research in Music Education, 52*(1), 16-28.

Owen, S. L. (2015). *Student perceptions of the efficacy of SmartMusic practice software* [Doctoral dissertation, California State University].

Ozeas, N. L. (1991). *The effect of the use of a computer assisted drill program on the aural skill development of students in beginning Solfege* (interval identification and sight singing) [Doctoral dissertation, University of Pittsburgh].

Paney, A. S., & Tharp, K. L. (2019). The effect of concurrent visual feedback on adult singing accuracy. *Psychology of Music*, 0305735619854534.

Perry, P. J. (2014). *The effect of flexible practice computer-assisted instruction and cognitive style on the development of music performance skills of high school instrumental students* [Doctoral dissertation, Shenandoah University].

Petty, C., & Henry, M. L. (2014). The Effects of Technology on the Sight-Reading Achievement of Beginning Choir Students. *Texas Music Education Research, 23*, 28.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception, 25*(2), 95-115.

Platte, J. D. (1981). *The effects of a microcomputer-assisted instructional program on the ability of college choral ensemble members to sing melodic configurations at sight* [Doctoral dissertation, Ball State University].

Potter, A. A. (2015). *Sight-singing systems in collegiate choral curricula: An examination of conductors' best practices at degree-granting institutions of the National Association of Schools of Music* [Doctoral dissertation, The Florida State University].

Price, H. E. (1992). Sequential patterns of music instruction and learning to use them. *Journal of Research in Music Education, 40*(1), 14-29.

Rutkowski, J., & Miller, M. S. (2003). The effect of teacher feedback and modeling on first graders' use of singing voice and developmental music aptitude. *Bulletin of the Council for Research in Music Education, 156*, 1-10.

Schmidt, C. P. (1995). Attributions of success, grade level, and gender as factors in choral students' perceptions of teacher feedback. *Journal of Research in Music Education, 43*(4), 313-329.

Skinner, B. F. (1965). *Science and human behavior.* Simon and Schuster.

Snider, M. T. (2007). *Choral sightsinging strategies in 1A, 3A, 4A, and 6A Kansas high schools: Instructional time allotted, teacher strategies, and materials used.* [Masters thesis, Emporia State University].

Steinhauer, K., & Grayhack, J. P. (2000). The role of knowledge of results in performance and learning of a voice motor task. *Journal of Voice, 14*(2), 137-145.

Timmers, C. F., Braber-Van Den Broek, J., & Van Den Berg, S. M. (2013). Motivational beliefs, student effort, and feedback behaviour in computer-based formative assessment. *Computers & Education, 60*(1), 25-31.

Thorndike, E. L. (1933). A proof of the law of effect. *Science, 77*, 173-175.

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology, 39*, 212-222.

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of educational research, 85*(4), 475-511.

Von Kampen, K. E. (2003). *An examination of factors influencing Nebraska high school choral directors' decisions to use sight-singing instruction* [Doctoral dissertation, The University of Nebraska – Lincoln].

Vispoel, W. P., & Austin, J. R. (1993). Constructive response to failure in music: The role of attribution feedback and classroom goal structure. *British Journal of Educational Psychology, 63*(1), 110-129.

Vujović, I., & Bogunović, B. (2012, July). Cognitive strategies in sight-singing. In *Proceedings Book of 12th International Conference on Music Perception and Cognition (ICMPC) and 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*, Thessaloniki, Greece.

Walls, K. C., Erwin, P. M., & Kuehne, J. M. (2013). Maintaining efficient ensemble rehearsals without sacrificing individual assessment: SmartMusic assessment could leave the director on the podium. *Journal of Technology in Music Learning, 5*(1), 4-16.

Welch, G. F. (1985). Variability of practice and knowledge of results as factors in learning to sing in tune. *Bulletin of the Council for Research in Music Education,* 238-247.

Wilson, P. H., Lee, K., Callaghan, J., & Thorpe, C. W. (2008). Learning to sing in tune: Does real-time visual feedback help? *Journal of Interdisciplinary Music Studies, 2.*

Appendix A

Pre/posttest melodies

Melody A

Melody B

Session melodies

Week 1

Week 2

Week 3

Week 4

Week 5

All melodies were selected from McGill, S., & Stevens Jr., H. M. (2003). 90 days to sight reading success: A singer's resource for competitive sightsinging. Houston, TX: AMC Publications.

Appendix B, Sight-singing Scoring

For the purposes of this study, every pitch and every rhythm will be scored separately. Therefore, every note will be worth two points.

Notes/Pitches

1 The melody be sung in the original key.
2. Participants may choose which octave they prefer.
3. Pitches may be within 50 cents of a half-step in either direction to be considered accurate.
4. Use of a solmization system is not scored- only pitch accuracy.
5. Pitches do not need to align with the click track to be considered accurate.
6. Participants may correct a single pitch by changing notes, sliding, or scooping, but may not go back once a subsequent pitch is attempted.
7. You may use clues to ascertain the participant's intended pitch including the click track, prior and following notes, and solfege syllables.
8. Accurate pitches are given a 1 (per note).
9. Inaccurate or omitted pitches are given a 0.

Rhythms

1. Rhythms must align with the click track and the quarter note indicator.
2. The pitch of a given note does not need to be accurate for the rhythm to be considered correct.
3. These are high school students and you are trying to score their ability, not their precision. Some leeway is appropriate.
4. Notes do not have to be performed for their full duration, but another note cannot be started before the current note duration is competed. Note pairs (two eighth-notes or a dotted quarter -eighth note pair) will likely need to be performed correctly to mark either correct in most situations.
5. The rhythm is considered incorrect if they change pitch or syllable during the note
6. Accurate rhythms are given a 1 (per note)
7. Inaccurate or omitted rhythms are given a 0

Other comments

1. In order for both pitches and rhythms to both be scored as accurate, the note and rhythm must be accurate.
2. When there is a discrepancy between pitches and rhythms, give preference to scoring the pitch as correct and mark the rhythm wrong.