Differences in error detection skills by band and choral preservice teachers

Laura A. Stambaugh
Differences in Error Detection Skills by Band and Choral Preservice Teachers

Laura A. Stambaugh

Abstract

Band and choral preservice teachers (N = 44) studied band and choral scores, listened to recordings of school ensembles, and identified errors in the recordings. Results indicated that preservice teachers identified significantly more errors when listening to recordings of their primary area (band majors listening to band, p = .045; choral majors listening to chorus, p = .012). Furthermore, band majors identified rhythm errors more frequently than pitch errors when listening to both types of ensembles, while chorus majors listened predominantly for pitch errors. A significant correlation existed between choral majors’ ability to detect pitch errors in choral recordings and their ability to detect pitch errors in band recordings (r = .556). Theory course grades were not significantly correlated to error detection abilities, but aural skills course grades were related to ability to identify pitch errors outside the primary area (band majors, r = .441; choral majors, r = .611).

Keywords

error detection, rehearsal, score study, conducting, preservice teachers

A music teacher’s ability to detect errors in ensemble performance is a critical skill for effective rehearsing and teaching. Cavitt (2003) found that almost 50% of band rehearsal time was spent on error correction, even as repertoire approached performance dates. Given the many kinds of possible errors (e.g., pitch, intonation, diction, and dynamics), compounded with classroom management demands, it is understandable that preservice teachers have difficulty identifying errors in the rehearsals they

1Georgia Southern University, Statesboro, GA, USA

Corresponding Author:
Laura A. Stambaugh, Box 8052, Music Department, Georgia Southern University, 1332 Southern Drive, Statesboro, GA 30460, USA.
Email: lstambaugh@georgiasouthern.edu
conduct. This difficulty may be further intensified when candidates conduct rehearsals in areas outside their primary content area (e.g., band majors rehearsing a chorus). Since many states issue teaching licenses that include all areas of music, it is quite possible that a preservice teacher may be hired to teach an ensemble outside his or her primary content area. While a respectable research base has examined instrumental preservice teachers’ abilities to detect errors in their primary performance area, it has neglected choral music education students and choral performances. Furthermore, previous research has not considered candidates’ error detection abilities outside their primary area. The purpose of this study was to examine band and choral majors’ abilities to identify errors in band and choral performances. It examined these abilities within and outside candidates’ primary performance area and also revisited the relationship between success in theory/aural skills courses and error detection ability.

Related Literature

At its core, the ability to identify errors in ensemble performance is an aural skill. Early research in this area examined if a candidate’s abilities in theory or aural skills classes were related to error detection ability. Results of previous studies have been split between majors. Instrumental majors showed no significant relationship between theory or aural skills classes and error detection ability (Brand & Burnsed, 1981; Sidnell, 1971), but a positive relationship was found for choral majors (Gonzo, 1971; Hansen, as cited in Taylor, 1963). While these studies occurred decades ago, they may still have relevance today. Davis (2010) noted that theory and aural skills courses are still being taught in a manner that is similar to decades ago. Researchers have found that other factors have also been unsuccessful in predicting error detection ability, including age, sex, number of years of study in the primary performance area, conducting experience, ensemble experience, and piano improvising ability (Brand & Burnsed, 1981; Hansen, as cited in Taylor, 1963). Teaching or conducting experience continues to be a topic of interest. Grunow (1980) found teaching and conducting experience did not improve error detection ability, but other studies have shown a positive relationship (Byo, 1997; Cavitt, 2003; Gonzo, 1971).

By manipulating the nature of the listening task, researchers have isolated several facets of error detection skill. The texture of the music affected instrumental candidates’ ability to detect errors, with candidates being more accurate when there were fewer parts in the texture (Byo, 1993, 1997) or in homophonic settings (Byo, 1997). Voicing interacted with texture, as errors in the soprano and alto ranges were more consistently identified than errors in tenor and bass ranges (Sheldon, 2004). Timbre was another significant factor. Instrumental majors identified more errors played on their primary instrument than on other instruments (VanderGheynst, 1978), although Swinehart (1994) and Sheldon were not able to replicate this result at the level of instrument families. The impact of tonality has only recently been investigated (Groulx, 2013), with chromatic errors being identified more accurately than diatonic errors. Overall, preservice teachers identified a low percentage of errors presented in error detection studies (48%: Byo, 1993; ~50%: Sheldon, 1998; 53.5%: Swinehart, 1994).
Some types of errors were identified more readily than other types of errors. When band students listened to reductions of full band scores, they identified more rhythm errors than pitch errors (Byo, 1993; Sheldon, 1998; Swinehart, 1994). Of course, a teacher needs to address errors while they are teaching on the podium and not merely in listening contexts. In an effort to bring more ecological validity to the task of error detection, Waggoner (2011) asked instrumental majors to detect errors in full band and sectional settings, both while conducting an ensemble and while simply listening to a recording. Candidates were more successful at identifying errors when just listening. By comparing full ensemble with sectional performances, Waggoner showed that pitch errors were most frequently identified in the full band context, while rhythm errors were most frequently identified in sectional contexts. Sheldon (2004) simulated the multiple-trial environment that could be present in an actual rehearsal. Given multiple opportunities to identify errors, wind majors continued to be most successful identifying pitch and rhythm errors, while also identifying articulation errors.

A considerable amount of effort has been put forth to determine how music teacher educators can best support the development of error detection skills in preservice teachers. While some research used the framework error detection ability, other relevant studies used the framework score study skills. Programmed instruction, both in-class or self-paced, has generally supported error detection ability for pitch and rhythm errors, and multiple-week programs were more beneficial than shorter programs (Constanza, 1971; Deal, 1985; Ramsey, 1979; Sidnell, 1971; exception, DeCarbo, 1982). A targeted approach to error detection skill through class discussion, ear training, or listening activities was found to be superior to conducting-only approaches for band and string performance (Sheldon, 1998; Stuart, 1979). When band majors learned to sing all the parts of a score, their focus on error detection shifted from rhythm errors at pretest to pitch errors at posttest (Byo & Sheldon, 2000). Score study at the piano has yielded inconsistent results (Crowe, 1996; Stwolinski, Faulconer, & Schwarzkopf, 1988). Listening to a correct aural referent while studying the score has supported error detection ability for listening tests (Crowe, 1996), but brief periods of listening with score study did not enable student conductors to identify errors (Van Oyen & Nierman, 1998).

Almost all previous research in error detection by preservice teachers has included instrumental students listening to instrumental recordings or ensembles. This bias highlights a pressing need to examine the error detection skills of choral preservice teachers listening to choral ensembles. Furthermore, it is not unusual for beginning teachers to teach outside their primary area (Andrews & Quinn, 2004). A saxophone major who has focused on preparing to teach band may find the difference between a .8 position and a full-time position is his ability and willingness to teach a choir class. Researchers have not investigated error detection skills for outside areas. Given these deficits in the research base, the purpose of this study was to examine the error detection skills of band and choral preservice teachers listening within and outside their primary area of emphasis. The following research questions were investigated:

**Research Question 1:** What types of errors do band and choral majors identify when listening to band and choral performances?
Research Question 2: What is the relationship between ability in theory/aural skills courses and ability to detect errors in performance?

Method

Participants

Forty-eight undergraduate music education students from one university in the Southeast United States participated in this study. Four candidates did not complete all the study protocols, which resulted in useable data sets for 15 choral majors (females, \( n = 10 \); males, \( n = 5 \)) and 29 band majors\(^1\) (females, \( n = 11 \); males, \( n = 18 \)). Participants were enrolled in a secondary music methods course, which is generally taken in the sixth semester of the degree program. All participants had completed a four-semester music theory and aural skills sequence, and two conducting courses. The informed consent process included granting permission for their theory and aural skills course grades to be collected from their transcripts.

Theory/Aural Skills Grades

The relevant music department courses were Theory I, II, III, and IV at three credits each and Sight-Singing/Ear Training I, II, III, and IV at one credit each. I collected these course grades from unofficial transcripts. The university grading system was letter grades only, with no plusses or minuses (A = 4.0, B = 3.0, C = 2.0, and D = 1.0). For the purposes of this study, I recoded the letter grades as numerical values: A = 95, B = 85, C = 75, D = 65, and F = 55.

Materials

I selected six band pieces (Grades 1–4) and six choral pieces (SAB [arrangement for sopranos, altos, and basses] and SATB [arrangement for sopranos, altos, tenors, and basses], a capella and accompanied) that commonly appear on state festival lists or in the Teaching Music Through Performance in Band/Choir series. I intentionally selected pieces including different keys, tempos, languages, and textures. Specific excerpts from the pieces were chosen from different sections (not always the beginning), and the mean length of excerpts was 35.5 seconds (range, 25–44 seconds). After entering the score excerpts into Finale, I saved audio files of the correct scores using MIDI (Musical Instrument Digital Interface) timbres. These aural referents were used during the score study process. MIDI timbres were used as the correct version of the audio scores because synthesized timbres have been used previously (Byo, 1993, 1997), and it is not unusual to find MIDI files available for score study.

Next, I added errors to individual parts in each score. The errors were not perfectly distributed among SATB voicings, but multiple errors were present in all band and chorus voices, including percussion. Pitch, rhythm, tempo, articulation, text, and dynamics/balance errors were based on mistakes that students were likely to make.
and had been identified in previous literature (Ramsey, 1979). It was possible for more than one error to be heard simultaneously (e.g., multiple instruments changed tempo, but one voice would not sing both a pitch and a rhythm error). I rehearsed and recorded these error-laden versions with a high school band and chorus. Ensemble recordings were used for the testing process because the tone quality and the timbre of developing groups is part of the complex task of error detection. The resulting performances included errors that I had planned, as well as errors that were not planned. Both intentional and unplanned errors were considered as viable errors to be detected in the study materials. After reviewing the recordings, I selected recordings of the three band and three choral pieces that had the most salient errors (see Table 1). Reliability was examined by two expert listeners from each area who also listened to the recordings and identified the errors they could hear. For the final study materials, the possible list of errors included only items that had been identified by at least two out of the three expert listeners (the third expert listener being myself).

### Procedure

Study sessions took place across four class meetings. The choral excerpts were studied and tested on the first 2 days, and the band excerpts were studied and tested on the third and fourth days. On the first day, candidates were given a printed copy of the correct version of the first choral excerpt. They were told,
You will hear a correct, MIDI version of the score today, but at the next class meeting you will hear a recording by a school ensemble who played/sang errors. At that time, you will need to circle and label on the score any errors you hear in pitch, rhythm, dynamics/balance, articulation, tempo, or text.

In addition, all preservice teachers received a printed copy of these directions. They were also told they could make any markings on the score that they wanted to during this preliminary score study, as the exact same copy of the score would be returned to them for the testing session. The correct recording was played, followed by 1 minute of silence for them to continue studying the score. This listening and silent study sequence was repeated two more times, for a total of three hearings. Next, preservice teachers repeated the entire listening and study process with the two remaining scores (score order was random). They then turned their scores in and were instructed not to seek out additional listening or score study opportunities with these scores before the next class.

On the second day, the copies of the score excerpts and the written directions were returned to individual participants. They were instructed to circle each error and label it as P = pitch, R = rhythm, A = articulation, D = dynamics/balance, T = tempo, or X = text. The first recording with errors was played one time, followed by 20 seconds of silence for participants to continue marking scores. The recording was then played a second and third time, with 20 seconds in between hearings. Next, this process was repeated with the second and third choral scores. On the third day, participants heard the correct, MIDI recordings of the band excerpts using the same protocol as the first day of the choral sessions. On the fourth day, participants completed the band error detection listening test, using the same protocol as the testing day of the choral session.

**Analysis**

Previous research has included theory and aural skills course grades, but it is not always clear exactly which course grades were used as dependent variables. Gonzo (1971), for example, used “composite” theory grades but did not clarify if that included course grades from two different courses or merely a composite of the activities occurring in one course. Because of the inconsistency in previous research, as well as the current paradigm of some programs teaching combined written and aural skills courses, I created one weighted theory and aural skills grade. The dependent variable of weighted theory score was determined by a weighted average. Since the university’s Theory grades were from a three-credit course but Sight-Singing/Ear Training grades were from a one credit course, Theory grades were entered into the mean equation three times.

Each participant’s score copies with circled and labeled errors were compared with the master list of errors. An error was considered correctly identified if it was circled and indicated as the correct type of error (pitch, rhythm, etc.). If a participant circled an error but did not indicate what kind of error was present, or indicated the wrong
type of error, the participant received no credit for that response. This level of stringency was decided a priori because analysis would already be complicated by having two different kinds of genres (band and chorus) plus two different areas of emphasis (within emphasis and outside emphasis) plus five different error possibilities (pitch, rhythm, dynamics, tempo, and articulation/text).

Results

A histogram of the mean weighted theory scores showed they met the assumption of normality. The overall mean weighted theory score was 87.62 \( (SD = 4.87) \). A t test for independent samples indicated no significant difference in the mean weighted theory score between band majors \( (M = 87.80, SD = 4.94) \) and choral majors \( (M = 87.26, SD = 4.91) \), \( t(42) = .735, p > .05 \). Therefore, preexisting ability in theory/aural skills courses did not serve as a confound in subsequent analyses.

The first research question concerned what types of errors band and choral majors could identify when listening to recordings of bands and choirs. Table 2 shows the mean number of errors correctly identified for each type of error. Band majors were most successful at locating rhythm errors, while choral majors were most successful at hearing pitch errors. A one-way analysis of variance indicated that band majors identified significantly more errors than choral majors for band performance, \( F(1, 37) = 4.319, p = .045, \eta^2 = .105 \). Likewise, choral majors identified significantly more errors than band majors for choral performances, \( F(1, 42) = 6.944, p = .012, \eta^2 = .142 \).

The second research question addressed the relationship between theory/aural skills grades and error detection skills. Pitch and rhythm errors were the only types of errors included in these analyses, due to the small number of items in the categories of articulation/text, dynamics, and tempo. Table 3 displays Pearson correlations among mean weighted theory scores, the total number of band and choral errors identified, and the number of rhythm and pitch errors identified. No significant relationship was found between mean weighted theory scores and ability to detect errors. However, when
aular skills grades were separated from theory course grades, significant correlations were found. For band majors, aural skills grades were significantly correlated with identifying chorus errors ($r = .446$, $p = .015$) and choral pitch errors ($r = .441$, $p = .017$). For choral majors, significant correlations were found between aural skills grades and identifying band pitch errors ($r = .611$, $p = .020$). In addition, choral majors demonstrated a significant relationship between the ability to detect pitch errors in choral and band recordings ($p < .05$).

### Discussion

Music teachers are often required to teach outside their primary area, but little is known about preservice teachers’ abilities in unfamiliar contexts. This study examined error detection ability within and outside preservice teachers’ primary area. When theory and aural skills course grades were combined in a weighted theory score, they were not related to error detection ability. Yet aural skills grades alone were related to the ability to identify pitch errors in ensembles outside preservice teachers’ primary area of emphasis. They identified significantly more errors within their primary area of emphasis than outside their emphasis. Band majors gave priority to rhythm errors in both listening contexts, while choral majors gave greater attention to pitch errors.

It is not surprising (and perhaps even affirming) that preservice teachers were most able to identify errors within their primary area of emphasis. This finding is consistent with VanderGheynst (1978), who determined band majors were most able to identify errors heard on their own instrument. Also similar to previous work (Gonzo, 1971; Hansen, as cited in Taylor, 1963), choral majors’ ability in aural skills courses was related to their ability to detect errors. Choral majors’ bias toward identifying pitch errors may be result of their extensive training in solfège. Conversely, band majors have previously been found to attend to rhythm errors more than pitch errors (Byo, 1993; Sheldon, 1998; Swinehart, 1994), except in a study that compared listening with single-instrument sections and with a full ensemble (Waggoner, 2011). Band majors’
emphasis on rhythm persisted even when they were provided with an aural referent during score study. In terms of ecological validity, the amount of score study time in research studies is quite limited when compared to what a teacher would normally do. Therefore, the attention to rhythm could be a visual artifact: It may be easier to see an incorrect rhythm during score scanning than to see an incorrect pitch. If so, this effect would be related to Sheldon’s (2004) suggestion that the visual layout of the staves on a score predisposes preservice teachers to attend to staves at the top of the page. If band majors’ bias to rhythm is not a visual artifact, it could be the result of their own previous experiences in band programs that emphasized rhythm. In addition, if a visual artifact was responsible for a preponderance of rhythm errors being identified, then that effect should have occurred for both the choral and the band majors.

It was interesting to find that each major gave more attention to one type of error across both listening contexts. Participants’ listening strategy, or focus of attention, remained constant even when the kind of ensemble differed. Due to their immersion in either the choral or the band world, preservice teachers may develop an error detection schema that reflects the biases of each world. Promisingly, the results of this study show transfer effects do exist for error detection: The ability for choral majors to identify pitch errors in choral settings was significantly correlated to their ability to identify pitch errors in band settings. Therefore, if preservice teachers could develop an additional schema for detecting errors in their outside area, transfer effects may enable them to improve skills in their primary area, as well.

**Limitations**

Several limitations lead to suggestions for further research. First, the theory and aural skills course grades assigned at this particular university were A (90–100), B (80–89), C (70–79), D (60–69), and F (<69). A more precise grading system such as a numerical system would enable future research to be more discriminating. Next, it was difficult to rehearse and record the school ensembles in the time allotted (1 hour). This constraint made it impossible to prepare all the planned errors in the scores. While such constraints are quite normal in school settings, future studies should plan for a rehearsal session and a recording session to increase the distribution of errors across all instrument and voice sections.

About participants, almost twice as many band majors as choral majors completed the study. Future research should endeavor to include more choral participants, especially given the dominance of band error detection studies. In addition, future participants should be queried as to their familiarity with any of the scores. It is quite possible some participants may have sung, played, or studied this repertoire in the past. Since familiarity with the repertoire may enhance error detection ability (Byo & Sheldon, 2000), future studies should consider this variable.

The score study and testing procedures were modified from prior investigations to facilitate continuity in this line of research. However, valuable insights could be gained by wandering further from these procedures. For example, if students did not hear the correct aural referent during score study, the test would reflect their error detection
ability after silent study alone. Finally, using recordings for error detection studies eliminates the spatial aspect of an ensemble. Live ensemble members are spread out in rows, enhancing auditory stream analysis (Bregman, 1990). While recordings allow for controlled stimuli and more efficient testing, they compromise ecological validity for the task teachers actually do. Future studies should continue to examine error detection in both recorded and live-listening contexts (Waggoner, 2011).

Implications for Music Teacher Education

A teacher’s ability to identify errors in ensemble performance is a complex skill with many component parts. Multiple studies across four decades have shown a limited relationship between theory/aural skills classes and error detection skills. This study refined that understanding by showing that aural skills courses support the ability to detect pitch errors but not rhythm errors. Unfortunately, ensembles are not limited to making pitch errors. Furthermore, a recent review of aural skills courses by Davis (2010) found that many aural skills courses were not teaching error detection from current perception and cognition viewpoints. Therefore, teacher education programs cannot rely on foundational courses such as theory and aural skills to be the lone scaffold for developing error detection skills. In addition to activities in conducting and methods courses, many individuals have developed successful programmed instruction or similar materials to support error detection skills. Some materials are now commercially available for individual or course use (Grunow, 1985; Jordan, 2006; Spradling, 2010).

Because band majors struggled to identify pitch errors in band performances, teacher educators might want to include more practice focusing on pitch. One effective strategy is for band majors to sing during score study (Byo & Sheldon, 2000). Band majors could also find value in studying choral scores and recordings, to develop their ability outside their primary area. Choral majors’ rhythmic accuracy scores were better in the choral-listening condition than in the band-listening condition. This may be because of the relationship between text and rhythm. Therefore, those responsible for teacher education programs might highlight this relationship between rhythm and text in choral settings. For band settings, choral majors will need to spend extra time preparing the rhythmic aspects of a band score.

The daily duties of many music teachers require them to teach outside their primary area of emphasis. Yet the demands of content-specific pedagogy within each area of emphasis make it difficult to include this level of breadth in preservice teacher degree programs. The results of this study allow teacher educators to target the specific error detection skills that need to be supported with band and choral majors. Having a more efficient approach might well make it possible to include breadth in our programs.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note
1. The uneven number of choral compared to band participants reflected student enrollment during data collection. The standard deviations for each group were similar for all dependent variables except detecting choral pitch errors (choral majors, SD = 2.957; band majors, SD = 1.719).

References


