Texas Music Education Research 2003

Reports of Research in Music Education
Presented at the Annual Meeting of the Texas Music Educators Association
San Antonio, Texas, February, 2003

Robert A. Duke, Chair TMEA Research Committee
School of Music, The University of Texas at Austin

Edited by: Edited by Mary Ellen Cavitt,
The University of Texas at San Antonio

Published by the Texas Music Educators Association, Austin, Texas

CONTENTS

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Keirsey Temperament and Musicians' Practice Preferences: An Exploratory Study

Nancy Barry and Russ Teweleit
University of Oklahoma

Effective practice is essential to the development of musicianship. Hallam cites the most applicable definition of practice found in the Oxford English Dictionary as being "a repeated exercise in an activity requiring the development of skill" (1997, p. 180). She argues, however, that effective practice is actually a multi-faceted activity that is much more complex than the above definition, taking a pragmatic approach by defining effective practice as "that which achieves the desired end-product, in as short of a time as possible, without interfering negatively with longer term goals" (1997, p. 181).

Presage factors are those characteristics and conditions existing prior to the current practice session and affecting the learning process. Learner characteristics such as level of expertise, approaches to practice, learning styles, motivation, self-esteem, and personality are considered presage factors (Biggs & Moore, 1993; Hallam, 1997). Assessment of these presage factors should help music educators better understand how their students learn. Such knowledge could then be used to develop more appropriate practice methods for individual students.

Research indicates that musicians approach practice in a variety of ways. In studying professional musicians who were multiple woodwind specialists, McLaughlin (1985) found that each of the fourteen musicians had her/his own method for transferring embouchure control, air support, tonguing, and vibrato production from one woodwind instrument to another. In another study, Hallam (1997) investigated the practicing of 22 professional musicians. Results revealed differences in orientation to practice, approach to detailed practice, and interpretation.

A number of studies have utilized the Myers-Briggs Type Indicator (MBTI) as a measure of individual differences in personality. Based upon Carl Jung's theory of psychological type, the purpose of the MBTI is "to identify, from self-report of easily recognized reactions, the basic preferences of people in regard to perception and judgment" (Myers & McCaulley, 1985, p. 1). The four different preference categories measured by the MBTI items are:

1. Extraversion/Introversion (EI). The Extravert focuses on the outer world of people and objects while the Introvert focuses on the inner world of concepts and ideas.
2. Sensing/Intuition (SN). Sensing types perceive the world through their senses while Intuitive types perceive through subconscious insight.
3. Thinking/Feeling (TF). Thinking types approach decisions objectively and systematically, contrasting with Feeling types who rely on subjective judgments and personal values.
4. Judging/Perceiving (JP). Judging types are organized and systematic whereas Perceiving types are flexible and spontaneous.

All possible combinations of these four pairs of opposites yield 16 different MBTI profiles.

Research also indicates significant relationships between individual differences such as personality type and many aspects of music behavior including listener's responses to music (Crickmore, 1968; Hedden, 1973; Lewis & Schmidt, 1991). Crickmore (1968) examined the

relationship between scores on the Maudsley Personality Inventory and a researcher developed test of music appreciation. Later, Hedden (1973) conducted a more comprehensive examination of listener responses to orchestral music finding that it was possible to identify similarities in reaction profiles of listeners. Lewis and Schmidt (1991) have since replicated Hedden's work with the purpose of reexamining the question with "specific focus on music listening responses as a function of the personality variables measured by the Myers-Briggs Type Indicator (MBTI)" (p. 313) thus eliminating the other variables.

Other investigations have studied successful music teachers (Kemp, 1982a; Wubbenhorst, 1994), applied teaching behaviors (Schmidt, 1989), gender differences (Kemp, 1982b), and practice attitudes (Barry, 2002). In a study that used the MBTI to examine whether a particular personality type occurs consistently among music teachers, Wubbenhorst found Extrovert/Introvert-Intuitive-Feeling-Judging (E/INFJ) to be the model-type among a sample of 31 music educators. Wubbenhorst's findings support earlier studies of university music majors, indicating that musicians tend to be either Extravert-Intuitive-Feeling-Perceiving (Henderson, 1984) or Extravert-Intuitive-Feeling-Judging (Rossman, 1979). Interestingly, music education students are more likely to be Perceiving while music educators tend to be more Judging (Wubbenhorst, 1994). A conclusive and comprehensive body of research concerning individual differences and music practice, however, does not yet exist.

Since the MBTI yields 16 different types, it can be difficult to obtain a sufficient sample size for statistical analysis (Barry, 2002). In contrast, the Keirsey Temperament sorter yields only four distinct personality types:

1. Rationals who are abstract in communicating and utilitarian in implementing goals. Rationals make up only five to seven percent of the general population.
2. Idealists who are abstract in communicating and cooperative in implementing goals. Idealists are relatively scarce and comprise only eight to ten percent of the population.
3. Artisans who are concrete in communicating and utilitarian in implementing goals. Artisans make up thirty-five to forty per cent of the population.
4. Guardians who are concrete in communicating and cooperative in implementing goals. Guardians comprise an even larger segment of the population than the Artisans. They make up at least forty to forty-five per cent of the population.

Music educators certainly acknowledge that effective practice is essential to the development of musicianship. Research indicates that different musicians approach practice differently (Hallam, 1997; McLaughlin, 1985), but a clear and consistent way to measure individual differences in relation to music practice has yet to be determined.

The purpose of this study was to investigate relationships between musicians' practice preferences, gender and Keirsey Temperament. A secondary purpose was to explore the feasibility of using the Keirsey Temperament Sorter as a measure of individual personality differences in relation to music practice preferences. This study addressed the following research questions:

1. What are musicians' stated practice preferences?
2. Do musicians' stated practice preferences differ according to university status (undergraduate, graduate, faculty)?
3. Do musicians' stated practice preferences differ according to principal instrument?
4. Do musicians' stated practice preferences differ according to gender?
5. Do musicians' stated practice preferences differ according to Keirsey Temperament?
6. Is the Keirsey Temperament Sorter an appropriate measure of individual personality in regard to music practice preferences?

Method

The data collection instruments used in this study were the Keirsey Temperment scale, a Music Practice Inventory (MPI), and an anonymous questionnaire requesting demographic information. The MPI was based upon an instrument that had been tested and validated in previous research studies (Barry, 2002). The MPI was used to ascertain participants' reactions to a series of 24 statements about practice on a Likert-type response scale ranging from 5 (Very Important) to 1 (Not Important). The 24 statements of the Music Practice Inventory appear in Table 1.

Music students and faculty at two universities in the southwestern United States were invited to participate on a voluntary basis. One hundred sixty-five musicians participated in the study (100 females, 65 males). The majority of the participants were undergraduate students (76.4%), with much smaller representation from graduate students (17.6%) and music faculty (6.1%). A level of $p<.05$ was employed for all tests of statistical significance.

Results

On the Keirsey test, the majority of the participants were Idealists (40.6%) or Guardians (37.0%) with much smaller numbers identified as Rationals (14.5%) or Artisans (7.9%). The distribution of Keirsey types among these musicians was in contrast to the general population distribution of 40-45% Guardians, 35-40% Artisans, 8-10% Idealists, and 5-7% Rationals.

On the Music Practice Inventory, statements rated highest in importance for effective music practice were "Making practice enjoyable," "Investing adequate time in practice," "Marking Music," "Beginning a new piece slowly and gradually increasing the tempo," "Engaging in self critique," and "Setting specific practice goals for each practice session." Lowest rated items were "Keeping a written record of practice time," "Keeping a journal of progress," "Keeping a written record of practice objectives," and "Following a written practice format" (See Table 1).

Table 1

Music Practice Inventory Response Means and Standard Deviations

	Item topic	All (*N*=165) Mean	SD	Rationals (*N*=24) Mean	SD	Idealists (*N*=67) Mean	SD	Artisans (*N*=13) Mean	SD	Guardians (*N*=61) Mean	SD
1.	Thinking about practice	3.97	.93	4.04	1.16	3.97	.91	4.10	1.14	3.93	.81
2.	Experimenting	4.01	.81	4.12	.94	4.03	.80	3.92	1.10	3.97	.71
3.	Following specific format	3.70	1.03	3.67	1.09	3.69	1.08	3.00	1.28	3.87	.85
4.	Mentally "going through"	3.85	1.03	3.87	1.01	3.76	.97	3.75	1.29	3.97	1.06
5.	Following written format	2.75	1.06	2.62	1.13	2.81	1.02	2.46	.88	2.79	1.13
6.	Reading articles/books	2.97	1.11	3.17	1.03	2.94	1.13	2.85	1.07	2.95	1.16
7.	Engaging in silent practice	3.52	.93	3.70	.86	3.58	.99	3.42	.90	3.39	.90
8.	Breaking practice up	3.98	1.07	4.12	.96	3.98	1.12	3.54	1.13	4.00	1.04
9.	Keeping practice journal	2.68	1.19	2.58	1.18	2.81	1.23	2.23	1.01	2.67	1.18
10.	Keeping written record	2.29	1.17	1.96	1.23	2.40	1.04	2.08	.95	2.31	1.30
11.	Analyze a new piece	3.63	1.12	3.71	1.08	3.56	1.08	3.38	1.56	3.73	1.09
12.	Focus on one section	3.77	1.01	3.42	1.31	3.79	.88	3.54	1.13	3.93	.95
13.	Setting specific goals	4.31	.80	4.58	.58	4.32	.81	3.92	1.11	4.28	.76

(Table 1 continues on next page.)

Table 1 (Continued)

Music Practice Inventory Response Means and Standard Deviations

	Item topic	All (*N*=165) Mean	SD	Rationals (*N*=24) Mean	SD	Idealists (*N*=67) Mean	SD	Artisans (*N*=13) Mean	SD	Guardians (*N*=61) Mean	SD
14.	Practice long as possible	2.96	1.18	2.62	1.17	2.96	1.11	2.61	1.45	3.18	1.18
15.	Practice with metronome	3.99	1.09	4.08	1.06	3.98	1.01	3.31	1.55	4.10	1.04
16.	Invest adequate time	4.44	.82	4.29	1.12	4.55	.56	3.77	1.54	4.51	.65
17.	Listen to recordings	4.21	.96	3.92	1.32	4.25	.97	4.15	1.07	4.29	.74
18.	Written practice objectives	2.70	1.17	2.42	1.32	2.87	1.15	2.46	1.20	2.69	1.12
19.	Marking music	4.39	.82	4.41	.71	4.45	.76	4.00	1.22	4.40	.82
20.	Audiotape practice	3.39	1.12	3.17	1.17	3.57	1.05	3.00	1.29	3.36	1.13
21.	Self critique	4.35	.84	4.50	.78	4.36	.79	4.23	1.17	4.29	.85
22.	Begin slowly	4.37	.84	4.30	.93	4.48	.72	4.08	1.19	4.34	.85
23.	Sight-read piece first	3.88	1.05	3.79	1.14	3.83	.95	4.08	1.11	3.92	1.12
24.	Make practice enjoyable	4.62	.66	4.46	.78	4.72	.49	4.62	1.12	4.57	.65

Note. 5 = Very Important, 4 = Important, 3 = Uncertain, 2 = Not Very Important, 1 = Not Important

Undergraduate students, graduate students and faculty generally agreed on most items. A MANOVA found significant differences for only two items: "Experimenting with different approaches to practice," with significantly higher ratings from graduate students than from undergraduate students, and "Following a specific practice format," with faculty ratings significantly lower than undergraduate student ratings.

Instrument family proved to be an important factor for several MPI items. A MANOVA procedure revealed nine significant differences among responses of musicians (See Table 2).

Table 2

Music Practice Inventory Response Means and Standard Deviations: Significant Differences by Instrument Family

Item topic	Brass (*N* = 41) Mean	SD	Strings (*N* = 15) Mean	SD	Percussion (*N* = 6) Mean	SD	Keyboard (*N* = 26) Mean	SD	Voice (*N* = 18) Mean	SD	Woodwinds (*N* = 59) Mean	SD
3. Following specific format	3.93	1.06	3.53	.99	4.33	.82	3.31	1.09	2.94	1.11	3.93	.81
4. Mentally "going through"	4.13	.91	3.93	1.03	4.17	.41	3.23	1.21	3.61	1.09	3.97	.95
5. Following written format	3.02	1.11	2.67	1.05	3.17	.98	2.19	.94	2.67	1.08	2.80	1.03
8. Breaking practice up	4.39	.89	4.27	.80	3.17	1.72	3.92	.94	3.83	1.15	3.78	1.12
9. Keeping practice journal	2.37	1.04	3.33	.98	3.83	1.17	2.11	1.03	2.94	1.43	2.78	1.16
10. Keeping written record	2.12	1.01	2.20	1.08	3.67	1.51	1.73	1.04	2.67	1.37	2.39	1.11
15. Practice with metronome	4.15	.91	4.20	.68	4.83	.41	3.23	1.18	2.67	1.19	4.47	.70
18. Written practice objectives	2.58	.97	3.47	1.25	3.67	1.03	2.08	1.13	2.56	1.10	2.81	1.18
22. Begin slowly	4.15	.99	4.13	.83	4.83	.41	4.36	.81	3.89	1.13	4.69	.50

Note. 5 = Very Important, 4 = Important, 3 = Uncertain, 2 = Not Very Important, 1 = Not Important

A MANOVA indicated significant main effects for Gender with males giving higher ratings for MPI scale items pertaining to "breaking practice up into two or more daily sessions," "practicing with the metronome," and "beginning a new piece slowly and gradually increasing the tempo."

Significant interactions between Gender and Keirsey type were found for several scale items. "Experimenting with different approaches to practice" received highest ratings from Artisan males (M = 4.5) and lowest rankings from Artisan females (M = 3.20); and "breaking practice up into two or more daily sessions" was rated highest by Guardian males (M = 4.39) and lowest by Artisan females (M = *2.60)*. "Practicing with the metronome" showed fairly consistent ratings between males and females within each Keirsey type with the exception of Artisans, with lowest ratings from female Artisans (M = 2.20) in comparison with male Artisans (M = 4.33). Male and female Rationals and Idealists reported almost identical ratings for "investing adequate time in practice" but Artisan females rated this item lower (M = 3.20) than Artisan males (M = 4.50) and Guardian males rated it lower than Guardian females. A similar trend was observed for "keeping a written record of practice objectives" with Guardian females giving lowest ratings (M = 1.60) in contrast with highest ratings from Guardian males (M = 3.17), on the other hand, Rational and Guardian females rated this item higher than their male counterparts. "Making audiotapes of practice" was ranked lowest by Artisan (M = 2.80) and Rational (M = 2.87) females, yet was ranked highest by Idealist females (M = 3.77) and Artisan males (M = 3.67).

The three MPI scale items rated most important by Rationals were “setting specific goals for each practice session” (M = 4.58), “making practice enjoyable” (M = 4.56), and “engaging in self critique” (M = 4.50). The three MPI scale items rated highest by Idealists were “making practice enjoyable” (M = 4.72), “investing adequate time in practice” (M = 4.55), and “beginning a new piece slowly and gradually increasing the tempo” (M = 4.48). Artisans expressed highest ratings for “making practice enjoyable” (M = 4.62), “beginning a new piece slowly and gradually increasing the tempo” (M = 4.23), and “listening to recordings of repertoire being practiced” (M = 4.15). Guardians reported highest importance for “making practice enjoyable” (M = 4.57), “investing adequate time in practice” (M = 4.51) and “marking music” (M = 4.39).

Significant main effects for Keirsey type were found for three items. "Setting specific goals for each practice session" received lowest ratings from Artisans (M = 3.72) in contrast with highest ratings from Rationals (M = 4.60). Guardians (M = 3.19) and Idealists (M = 2.98) reported higher importance for "Practicing as long as possible during each session" in contrast with Rationals (M = 2.64) and Artisans (M =2.18). "Investing adequate time in practice" received relatively high ratings from all participants, with Idealists (M = 4.57) and Guardians (M = 4.50) reporting greater importance than Rationals (M = 4.27) and Artisans (M = 3.91) (See Table 1).

Discussion and Implications

What are musicians' stated practice preferences?

The results of the Music Practice Inventory indicated a high level of agreement on both “Making practice enjoyable” (M = 4.62), and “Investing adequate time in practice” (M = 4.44) across all Keirsey Temperment qualities. While these results are not surprising, they do serve as a reminder that teachers should make efforts to help their students find ways to make practice sessions enjoyable. Perhaps more importantly, applied instructors can help their students in setting up a practice schedule that allows the student to invest adequate time in practice.

Do musicians' stated practice preferences differ according to university status (undergraduate, graduate, faculty)?

Undergraduate students, graduate students, and faculty generally agreed on most items. This level of agreement is predictable, as it is likely that music students would echo their professors' attitudes about practice. The only exceptions were faculty and graduate students expressing higher importance for "experimenting with different approaches to practice" and lower importance for "following a specific practice format" than undergraduates. It seems, then, that more experienced musicians value a more varied, experimental approach to practice. However, since the majority of

participants in this study were undergraduate students, these results are inconclusive. More research is needed with larger numbers of faculty and graduate students participating. Subsequent research could reveal important information regarding differences in the way that student musicians and more advanced musicians approach practice.

Do musicians' stated practice preferences differ according to principal instrument?

The results of the MPI suggest that instrumentalists and more specifically wind players (M = 3.93) place a higher value on "Following a specific format" than vocalists (M = 2.94). This suggests that structured practice sessions may have greater appeal to instrumentalists. Perhaps this preference is due to the technical and mechanical elements involved in playing an instrument. In any event, it seems appropriate to adjust the practice regime according to the principal instrument/voice.

It is interesting that MANOVA procedures revealed significant differences between brass players' and other instrumentalists' responses to "Mentally going through a new piece before playing it" (M = 3.93) and "Breaking up practice into two or more daily sessions" (M = 3.93). These preferences may be due to the physical nature of brass playing and brass players' need to rest periodically. However, knowing this could be helpful to music teachers. This would be especially true in situations where non-brass musicians are involved in the instruction of brass players. However, due to the small number of participants representing some instrument families, such as percussion, these results are inconclusive. Additional research with larger representation from all instrument families is needed.

Do musicians' stated practice preferences differ according to gender?

A number of gender effects were observed in this study, with males giving higher ratings for MPI scale items pertaining to "breaking practice up into two or more daily sessions," "practicing with the metronome," and "beginning a new piece slowly and gradually increasing the tempo."

Several interesting interactions between Gender and Keirsey type were also observed. For example, Artisan females gave the lowest rankings for "Experimenting with different approaches to practice," "Breaking practice up into two or more daily sessions," "Practicing with the metronome," and "Making audiotapes of practice." Artisan females also generally disagreed with Artisan males. In contrast, Artisan males gave the highest rankings for "Experimenting with different approaches to practice," "Practicing with the metronome," and "Making audiotapes of practice." It certainly seems that males and females have different attitudes about some aspects of practice, even within the same Keirsey temperament type. While these contrasts are interesting and worthy of further investigation, results of the present study are inconclusive due to the small number of Artisans (N = 13) participating.

On the other hand, Guardians (N = 64) comprised a large percentage of the participants in this study. The significant differences found between Guardian males and females in "keeping a written record of practice objectives" with Guardian females giving lowest ratings (M = 1.60) in contrast with highest ratings from Guardian males (M = 3.17) suggest that different genders may place different values upon these practice activities. Again, the implication seems to be that a "one size fits all" approach to teaching music practice is not the most appropriate. Music teacher might be advised to consider both temperament and gender when helping students design a practice regime.

Do musicians' stated practice preferences differ according to Keirsey Temperament?

Results of this study suggest that some practice preferences differ according to Keirsey Temperament. For instance, an artisan would be much less likely to set specific goals for a practice session than a Rational. An Idealist or Guardian would be more likely to practice as long as possible whereas a Rational or an Artisan would not. With further study, this information could prove useful in helping teachers guide their students toward developing a more effective practice plan. The Keirsey Temperament Sorter could be used to help develop an individual "practice prescription" for each student.

Is the Keirsey Temperament Sorter an appropriate measure of individual personality preferences in regard to music practice preferences?

This study indicates that some practice attitudes vary according to Keirsey Temperament and gender. This research reinforces the importance of individualizing music instruction and working with each student to develop an optimal practice regime. Results of this study are not conclusive, but findings suggest that the Keirsey Temperament Sorter may be a useful tool for determining some types of individual differences among musicians. Additional research is needed to explore these implications and to develop research-based recommendations for better accommodating individual differences among music students.

References

Barry, N.H. (2002, April). *Practice and personality: An exploratory study of the relationships among musicians' practice preferences, gender and Jung-Myers-Briggs Personality type*. Paper presented at the meeting of the Music Educators National Conference, Nashville, TN.

Biggs, J. & Moore, P.J. (1993). *The process of learning*, New York: Prentice Hall.

Crickmore, L. (1968). An approach to the measurement of music appreciation (II). *Journal of Research in Music Education, 16*, 291-301.

Hallam, S. (1997). Approaches to instrumental music practice of experts and novices: Implications for education. In H. Jorgensen & A.C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 179-231). Oslo, Norway: Norges Musikkhogskole.

Hedden, S.K. (1973). Listeners' responses to music in relation to autochthonous and experiential factors. *Journal of Research in Music Education, 21*,225-238.

Kemp, A.E. (1982a). Personality traits of successful student music teachers. Proceedings of the Ninth International Research Seminar. *Psychology of Music*, Special issue, 72-75.

Kemp, A.E. (1982b). The personality structure of the musician. III. The significance of sex differences. *Psychology of Music*, 10(1), 48-58.

Lewis, B.E., & Schmidt, C.P. (1991). Listeners' response to music as a function of personality type. *Journal of Research in Music Education, 35*(4), 311-321.

McLaughlin, D.B. (1985). *An investigation of performance problems confronted by multiple woodwind specialists.* (Ed.D. thesis, Columbia University Teachers College, 1985). *Dissertation Abstracts International, 46,* No. 09A, 2610.

Myers, I. B. and McCaulley. M. H. (1985*). Manual: A Guide to the development and use of the Myers-Briggs Type Indicator, Palo Alto, CA: Consulting Psychologists Press.*

Rossman, R.L. (1979, October). *MBTI types of music education students: Morningside College, Sioux City, IA.* Paper presented at the Third National Conference on the Myers-Briggs Type Indicator, Philadelphia, PA.

Schmidt, C.P. (1989). Applied music teaching behavior as a function of selected personality variables. *Journal of Research in Music Education, 37*(4), 258-271.

Wubbenhorst, T.M. (1994). Personality characteristics of music educators and performers. *Psychology of Music, 22*(1), 63-74.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

The Effects of Focused Observation on Novice Teachers' Self-Evaluation of Piano Instruction

Cynthia Benson
Bowling Green State University

Effective teaching is defined as either that which causes students to learn and grow or that which is accepted by teachers and other educational professionals. The definition and measurement of effective teaching are difficult to separate since determining the first seems to predetermine the resolution of the second. One reason for defining and measuring effective teaching is to improve teaching skills (Tuckman, 1995).

The student teaching or internship experience is generally regarded as an integral and required part of the teacher preparation experience. It provides an opportunity for novice teachers to assess and define effective teaching skills. Directed student teaching is recommended for an effective piano pedagogy program. In fact, it has been suggested that the more opportunity for participation in teaching experiences, the stronger the program (Uszler & Larimer, 1984). Observation and intern teaching are indispensable training tools that are essential in the development of real teaching skills (Goss, 1997). A survey conducted by the Committee on Intern Teaching of the National Conference on Piano Pedagogy in 1990 found that practice teaching was required in more programs than was found by a survey conducted four years earlier. However, a report from the Seminar on Observation and Intern Teaching addressed the lack of observation intern teaching opportunities in piano pedagogy (Goss, 1997). Although peer teaching experiences may occur in piano pedagogy courses, field-based experiences have been found more productive for and viewed more positively by novices than are peer teaching (Duke 1999/2000).

Comparisons of expert and novice teachers have shown differences in their knowledge structure or schemata. Novice teachers' schemata seem to be less elaborate, less interconnected and accessible than those of experts, while experts have more domain-specific (known facts) and procedural (processes and conditions) knowledge (Graham, French & Woods, 1993). Researchers have inferred that the cognitive schemata of experts are more complex, interconnected, and accessible (Butler, 2001). Goolsby (1999, 1997, and 1996) compared novice and expert teachers preparing band compositions, providing verbal instruction, and selecting their use of rehearsal time. The results of these studies indicate that novice and expert teachers differ in attention to overall sound, efficiency of verbal instruction, and time spent in performance. Duke (1999/2000) reports that expert teachers' intervals of verbalizations and modeling are briefer than are those of less expert teachers and novices. Experts also seem to address different performance variables than are addressed by less expert teachers as well as employ more repetition and drill in rehearsal.

Studies have also investigated the use of self-evaluation for development and improvement of teaching skills. Reflective thinking has been viewed as (1) an essential part of developing complex understandings of effective professional practice; (2) a process that makes meaning of the learning experience; (3) a type of thinking that is closely related to experience; and (4) a set of skills including self-awareness, description, critical analysis, synthesis, and evaluation (Fallon & Brown, 2002). One purpose for developing the ability to reflect is that learning about practice is maximized through taking responsibility for one's own professional development (Walkington, et al, 2001). Kusnic & Finley (1993) state that student self-evaluation is one strategy for developing

skills of self-reflection, bringing students more fully into the education process, and helping them build active and meaningful relationships with the material they are studying. Self-evaluation helps students make meaning, derive relevance, and build coherence through their education experiences. The self-viewing experience has attained wide popularity as a method for inducing self-awareness and control and developing interpersonal understanding and skill.

It has been suggested by Hoy & Miskel (1987) that self-evaluation, if properly structured, can foster greater involvement of students in the teaching/learning process, heightening student enthusiasm and motivation to perform well. Previous studies have shown successful use of self-analysis to improve music teacher performance (Benson, 1989; Rosenthal, 1985), nonmusic majors skills in teaching music (Killian, 1981; Prickett, 1987), and music therapy skills (Alley, 1980; Decuir & Jacobs, 1990).

The use of videotape recordings along with self-analysis to improve skills has also been investigated. Berger (1978) found unanimity among investigators who utilize videotape confrontation that through this experience insight can be heightened producing marked changes in behavior. The combined use of self-analysis and videotape recordings has been found to substantially increase defined skills of music therapy students (Alley, 1980, 1982) and conducting students (Johnston, 1993; Yarborough, 1987; Yarborough, Wapnick, & Kelly, 1979). Price (1992) examined the effects of instruction, teaching practicum, feedback from instructor and videotaped self-observation on undergraduates' use of complete sequential patterns (teacher/student interaction sequence of teacher presentation of task, student interaction with task and teacher, and teacher specific feedback). Results indicated that the use of videotape self-observation along with clear operational definitions trained students to complete sequential patterns and provide better feedback. This was especially noteworthy where no course instructor feedback was used suggesting that students became more independent in assessing and improving teaching by relying less on external feedback.

Broyles (1997) investigated the use of videotape self-analysis on role development of student teachers in music. The results from questionnaires distributed to the student teachers, cooperating teachers and university supervisors indicated that students felt that viewing videotaped examples of their teaching was a helpful procedure in their learning to teach. The cooperating teachers reported that the videotape analysis seemed to help their student teachers improve in their teaching and become more aware of how well their pupils were learning. University supervisors' responses revealed that videotape analysis strengthened participants' teacher identity, increased their commitment to refining teaching tasks and skills, and enhanced their concern for pupil learning.

Comparisons of supervisor and student-teacher self-evaluations show that student teachers tend to give themselves higher ratings (Briggs, Richardson, & Sefzik, 1985; Cassidy, 1990, Colwell, 1995). A high degree of agreement, however, was found between ratings made by music students and their supervisors in terms of perceptions of behaviors during an undergraduate practicum (Decuir and Jacobs, 1990).

The effectiveness of feedback obtained through systematic self-observation was also found to be undifferentiated from feedback given by course instructors. A consistent finding across a number of investigations is that teachers who obtain a more clear and precise picture of their teaching behavior are likely to make positive changes in their teaching, through systematic self-observation (Duke, 1999/2000).

It is critical that students learn to view their own teaching accurately. This may be accomplished through focused observation of their teaching, as focusing attention can affect evaluation of instruction. Observation focus has been investigated by comparing undergraduate and graduate music education/therapy majors (Standley and Greenfield, 1987) and trained and untrained observers (Duke, 1987). Yarborough & Hendley (1990) found that ratings for each of the 10 categories of choral rehearsal excerpts were higher for the teacher observation focus than for the student observation focus, and more comments were made about the teacher than the students regardless of the focus of observation. Focus on teacher behavior allows subjects to perceive a greater frequency of disapproval feedback and less positive affect than actually was observed (Duke & Prickett, 1987).

Cassidy (1993) explored the use of delivery and instruction forms for systematic self-analysis of teaching videotapes. In certain categories, focusing subjects' observation through the use of

specific forms was found to be helpful: subjects became more reliable with and deviated less from an experienced observer. Although subjects in the study identified more instances of good teaching than the experienced observer, practice in using the forms across four teaching segments decreased this discrepancy. The purpose of this study was to determine the effects of focused observation and videotape self-observation on novice teachers' self-evaluation of piano instruction.

Method

Participants in this study were graduate (6) and undergraduate (1) students enrolled in a piano pedagogy practicum. During 12 weeks, 5 participants taught one 30-minute studio (one-on-one) lesson each week while 2 taught a 30-minute group piano class. Ages of the students of the participants ranged from 5.5 to 9.5 and each had received less than one year of study previous to taking lessons from the participants.

Following each lesson, participants completed a self-evaluation form on which they described: (1) what aspects of student performance were addressed; (2) student improvement or no improvement in regards to these aspects; (3) targeted teacher and student behaviors to be improved and strategies to be incorporated into the next lesson.

Participants' piano instruction was videotaped every 3 weeks for a total of 4 lessons. With each lesson that was videotaped, participants completed the self-evaluation form described above before watching the videotape. Upon viewing the videotape, another form (video observation) was completed. After the first videotape, the video observation form (#1) requested the exact same information as the self-evaluation form described above. For the second videotape viewing, the video observation form (#2) asked the participants to indicate the teacher and student behaviors observed as well as successful and unsuccessful performance for each student activity. For the third videotaping (video observation form #3), participants, using the SCRIBE computer observation program, observed frequency, total time, percentage of lesson time, mean duration, and rate per minute of teacher talk and performance, student talk and performance, and teacher verbalizations (information, directives, specific feedback, and nonspecific feedback, and off-task verbal) (see Appendix A for all observation forms). Upon the fourth videotaping (video observation form #4) participants were to choose a 12-15 minute segment in which they were actively engaged in improving some aspect of student behavior. Within this segment, participants were to: (1) list teacher instructions for each task; (2) monitor student successful/unsuccessful activities, teacher response to successful/unsuccessful trials; (3) goals met in instructional sequence; and (4) indicate instructions for student practice.

Each of these video observation forms also asked participants to target teacher and student behaviors that they wanted to improve and the strategies to be incorporated to improve those behaviors. Participants met with the practicum supervisor regularly, but the supervisor did not give feedback regarding the participants' teaching until after they had completed the forms.

The comparison of the responses given by the participants on the self-evaluation forms and the video observation forms allowed for the study of the effects of: (1) video feedback and focused attention on participants' perception of student performance aspects addressed in lesson; (2) video feedback and focused attention on participants' perception on student improvement or no improvement on performance aspects; (3) video feedback and focused attention on participants' targeted teach and student behaviors for improvement; (4) video feedback and focused attention on participants' strategies to improve targeted student and teacher behaviors.

Results

Comments (N=982) written by the participants were categorized according to: (1) what aspects of student performance were addressed; (2) student improvement or no improvement in regards to these aspects; and (3) targeted teacher and student behaviors to be improved and strategies to be incorporated into the next lesson.

Aspects of Student Performance Addressed

Participants responded to aspects of student performance that were addressed in each lesson (see Table 1). Technique was listed most frequently (*n*=95) followed by reading (*n*=77) and rhythm (*n*=65). The participants were working with beginning piano students, so these aspects along with keyboard topography (*n*=25) would logically be addressed frequently over these 12 lessons. The table also shows that participants addressed fingering only until Lesson #6, whereas dynamics was addressed starting during the third lesson. The number of comments ranged from eighteen (Lesson 12) to fifty-five (Lesson 2 video). The fewer aspects addressed in the last lesson could indicate that more attention was given to a smaller number of performance aspects. The most number of comments occurred using video observation #1 in which participants' attention was focused on the same questions as the anecdotal logs.

The participants responded with comments regarding nonmusical as well as musical aspects. Comments regarding teacher behavior (presentation of material, pacing, planning) were most frequent (*n*=86) followed by evaluation of student performance (*n*=55), self-evaluation (*n*=14), and student behavior (*n*=10). The participants' focus of attention seemed to be on their own behavior as much as that of their students' even though they were to respond only to student performance aspects of the lesson.

> *"…I thought my teaching went very well…my lesson plan in terms of time management was under control…"*

> *"…I tried everything, right down to playing on top of his fingers so he could feel what it is like when I do it…"*

The participants commented on more teacher behaviors after watching the first and second videos, but not after the third and fourth videos. Perhaps it was due to the participants' attention focus on observation forms #3 and #4 as the forms called for the participants to comment on student success within the sequence of presentation and student play and talk during the lesson.

Table 1
Number of Student Comments on Aspects of Student Performance Addressed in Each Lesson

	Lesson/Anecdotal Log #														
Musical	1	2	V	3	4	5	V	6	7	8	9	10	11	12	Total
Technique	13	9	10	7	7	7	9	3	9	5	2	5	7	2	95
Reading	4	4	3	5	6	8	6	6	4	10	5	6	7	3	77
Rhythm	4	6	6	3	5	4	8	4	5	5	4	2	4	5	65
Topography	5	5	4	3	1	0	3	1	2	0	1	0	0	0	25
Finger #'s	2	1	2	3	0	0	1	0	0	0	0	0	0	0	9
Duet	1	0	0	0	0	1	0	1	0	0	2	0	0	1	6
Creativity	1	0	1	0	0	0	1	1	0	0	0	0	2	0	6
Ear training	1	0	0	0	1	0	1	3	1	0	0	0	0	0	7
Review	0	1	4	1	3	0	0	1	2	0	2	0	1	0	15
Written work	0	1	1	0	1	1	0	0	1	0	0	1	0	0	6
Dynamics	0	0	0	3	2	0	3	2	1	0	2	1	2	2	18
Nonmusical															
Eval. of St.	2	1	4	4	2	5	0	4	11	11	4	2	1	4	55
S. behavior	0	0	2	2	0	1	0	0	0	1	2	0	2	0	10
T. behavior	0	6	14	8	4	11	0	7	8	11	10	0	6	1	86
T. self-eval.	0	0	4	2	2	1	0	3	0	1	0	0	1	0	14
Other	1	0	0	0	1	1	0	0	0	0	0	0	0	0	3
Total	34	34	55	41	35	40	32	36	44	44	34	17	33	18	497

Student Improvement or No Improvement in Regards to These Aspects

As seen in Table 2, participants commented most on student musical aspects with respect to improvement or no improvement (*n*=107). The most number of statements regarding these aspects were made after watching the first video (*n*=19). Again, this could have been due to participants' focus of attention on these questions. Comments regarding student nonmusical behaviors (*n*=22) and teacher behaviors (*n*=6) were less frequent. There were more positive comments (*n*=120) than negative comments (*n*=47) regarding improvement of behaviors.

A subcategory of the response to this question was in regards to what the participants attributed improvement or no improvement. These comments attributed the teacher most frequently for student progress or lack thereof (*n*=45).

"…Student improved on all aspects except the strong finger. I believe I was not clear with my directions on this topic…"

"…I talked too fast and introduced too many things…"

"…the student benefited when I broke the piece down into smaller units and made more rapid progress…"

Fewer comments made reference to the student (*n*=23), amount of time (*n*=6), home or environment outside of the lesson (*n*=6), or the difficulty of the skill (*n*=2) for improvement or no improvement of student performance aspects.

"…the students were quite cooperative and involved and that made my teaching a lot happier…"

"…he started to understand my requirements about right hand position and fingers, but improving takes time…"

The largest number of comments by the participants attributing improvement or no improvement to the teacher could also be due to the focus of attention on the teacher. The second comment above is indicative of the comments attributing improvement or no improvement to the students. While participants attributed progress/no progress directly to teacher behaviors (talk too much, breaking material down into smaller sections), the students' social behaviors (attentive, cooperative, liked piece, was enthusiastic, enjoy) were given credit for success or no success more often than were students' academic behavior.

Table 2
Number of Participant Comments on Improvement or No Improvement

		Lesson/Anecdotal Log #													
Aspects		1	2	V	3	4	5	6	7	8	9	10	11	12	Total
	S. Musical	6	5	19	10	8	7	8	12	9	4	8	7	4	107
	S. Nonmusical	5	4	3	2	0	0	3	3	1	0	0	1	0	22
	T. Behavior	1	1	3	0	0	0	0	0	1	0	0	0	0	6
Improvement															
	Positive	7	12	15	10	9	13	12	11	7	4	8	7	5	120
	Negative	5	5	6	2	8	3	4	4	1	2	2	3	2	47
Attribution															
	Teacher	3	3	0	4	6	6	7	5	1	2	5	3	0	45
	Student	3	0	0	5	3	2	2	1	1	0	3	2	1	23
	Home/beyond lesson	2	1	0	0	0	0	0	0	0	0	1	0	2	6
	Subject	1	0	0	0	0	0	0	0	0	0	0	1	0	2
	Time	2	0	0	0	0	0	0	1	1	1	0	0	1	6
	Total	35	31	46	33	34	31	36	37	22	13	27	24	15	384

Targeted Teacher and Student Behaviors to Be Improved and Strategies to Be Incorporated into the Next Lesson

Participants targeted student musical behaviors (*n*=93) more often than nonmusical behaviors (*n*=16) or teacher behaviors (*n*=26) for improvement (see Table 3). Strategies prescribed by the participants called for changes in teacher behaviors (*n*=34), addition of different activities (*n*=27), use of different material (*n*=8), change in student behaviors (*n*=2) and other (*n*=7). Upon viewing the first video, more comments targeted teacher behavior for improvement (*n*=9) than before viewing the tape (*n*=6). The participants seemed equally concerned with *how* to improve targeted behaviors and *with what* (material and activity).

"…improve by having more movement activities…"

"…teacher needs to be a little slower in explanations….:"

"…the next time with wrist problem, I will have him play something he knows from memory and concentrate solely on his RH wrist, watching it to make sure that it doesn't drop, matching it to his correct left hand…"

"…I would like to bring in rhythm flash cards and ask her to compose a song or rhythm with steps…"

"…introducing in small steps…"

Table 3
Number of Participant Comments on Targeted Teacher and Student Behaviors to Be Improved and Strategies to Be Incorporated into the Next Lesson

		Lesson/Anecdotal Log #													
		1	2	V	3	4	5	6	7	8	9	10	11	12	Total
Target															
	S. Musical	8	6	2	8	9	8	4	9	8	10	7	8	6	93
	S. Nonmusical	6	3	2	0	2	2	0	0	0	0	0	1	0	16
	Teacher	0	6	9	2	0	5	0	1	1	2	0	0	0	26
Strategies															
	Teacher	3	3	2	0	3	3	3	2	6	1	6	2	0	34
	Student	1	0	0	0	0	1	0	0	0	0	0	0	0	2
	Material	4	0	0	0	1	1	1	0	0	0	1	0	0	8
	Activity	1	4	2	4	2	2	2	2	2	3	0	0	3	27
	Other	1	4	0	0	0	0	2	0	0	0	0	0	0	7
	Total	24	26	17	14	17	22	12	14	17	16	14	11	9	213

Observation #3 included participants observation of teacher talk, teacher play, student talk, student play through the use of the computer observation program, SCRIBE. The mean percentage of teacher talk was 42.2; teacher play was 4.9; student talk was 19.26; and student play was 26.6. Although participants focused their attention on these aspects, there were no comments regarding them in relation to targeted behaviors or strategies.

Observation #4 asked participants to respond to a question regarding any deviations from the lesson plan. Comments made were with respect to deviating due to more time left in the lesson, not enough time in the lesson, and moving backwards within a sequence of presentation due to student performance.

Also noted were participant statements regarding why they made a decision before, during or after a lesson. For half of the 12 comments, participants noted student academic behaviors as the reason for their decision while student social behaviors were cited as reasons for their judgments.

From the results, participants seemed to focus on teacher behavior although not as much as student academic behavior. They commented more on improvement/no improvement of student musical behavior and more positively than negatively in regards to improvement/no improvement.

Participants also attributed the teacher more frequently for this improvement/no improvement. Student music behaviors were targeted for improvement more than nonmusical behavior or teacher behaviors while participants' strategies called for a change in teacher behaviors more often than a change in student behavior.

References

Alley, Jayne M. (1980). The effect of self-analysis of videotapes on selected competencies of music therapy majors. *Journal of Music Therapy, XVII* (3), 113-132.

Alley, Jayne M. (1982). The effect of videotape analysis on music therapy competencies: An observation of simulated and clinical activities. *Journal of Music Therapy, 19,* 141-160.

Benson, W. L. (1989). The effect of models, self-observation, and evaluation on the modification of specified teaching behaviors of an applied music teacher. *Update, 7*(2), 28-31.

Broyles, J.W. (1997). *Effects of videotape analysis on role development of student teachers in music*. Unpublished doctoral dissertation, The University of Oklahoma.

Bruce, W. (1995). Assessing effective teaching. *Peabody Journal of Education, 70*(2), 127-139.

Butler, A. (2001). Preservice music teachers' conceptions of teaching effectiveness, microteaching experiences, and teaching performance. *Journal of Research in Music Education, 49*(3), 258-272.

Cassidy, J.W. (1993). A comparison between students' self-observation and instructor observation of teacher intensity behaviors. *Bulletin of the Council for Research in Music Education, 115,* 15-29.

Colwell, C.M. (1995). Effect of teaching setting and self-evaluation on teacher intensity behaviors. *Journal of Research in Music Education, 43* (1), 6-21.

Duke, R. A. (1999/2000). Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education, 143,* 1-48.

Duke, R.A. (1987). Observation of applied music instruction: The perceptions of trained and untrained observers. In *Applications of Research in Music Behavior,* Ed. Clifford Madsen and Carol Prickett, pp. 115-124.

Duke, R.A. & Prickett, C. A. (1987). The effect of differentially focused observation on evaluation of instruction. *Journal of Research in Music Education, 35*(1), 27-37.

Goolsby, T.W. (1996). Time use in instrumental rehearsals: A comparison of experienced, novice, and student teachers. *Journal of Research in Music Education, 44*(4), 286-303.

Goolsby, T. W. (1997). Verbal instruction in instrumental rehearsals: A comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45*(1), 21-40.

Goolsby, T. W. (1999). A comparison of expert and novice music teachers' preparing identical band compositions: an operational replication. *Journal of Research in Music Education, 47*(2), 174-187.

Goss, L. (1997). Pedagogy Saturday session summaries: Observation and intern teaching. *American Music Teacher, 47*(1), 40-41.

Graham, K.C., French, K.E., & Woods, A.M. (1993). Observing and interpreting teaching-learning processes: Novice PETE students, experience PETE students, and expert teacher educators. *Journal of Teaching in Physical Education, 13,* 46-61.

Hoy, W.K. and Miskel, C.G. (1987). *Educational administration: Theory, practice, and research.* (3rd ed.). New York: Random House.

Johnston, Hugh. (1993). The use of video self-assessment, peer-assessment, and instructor feedback in evaluating conducting skills in music student teachers. *British Journal of Music Education, 10,* 57-63.

Killian, J.K. (1981). Effect of instruction and feedback on music teaching skills. *Journal of Music Therapy, 15,* 15-20.

Lethco, L.M. (1999). *Preparing undergraduate music majors to teach beginning instrumentalists: The effects of self-evaluation, teacher observation, and performance-oriented instructional approaches on teacher behaviors and pupil responses.* Unpublished doctoral dissertation, The Louisiana State University.

Murray, K. & Williams, K. (1995). Intern teaching: Summary of statistics from the 1990 questionnaire. In Richard Chronister and Patrick Meader (Eds.), *The National Conference on Piano Pedagogy: Proceedings and Reference*. Los Angeles, California: The National Conference on Piano Pedagogy, 202-205.

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

Prickett, C. A, (1987). The effect of self-monitoring on the rate of a verbal mannerism of song leaders. In *Applications of Research in Music Behavior,* Ed. Clifford Madsen and Carol Prickett, pp. 125-134.

Standley, J.M. & Greenfield, D. G. (1987). The effect of a focused observation task and its transfer on the analysis of a kindergarten music class by pre-senior versus pre-internship music education/therapy majors. In *Applications of Research in Music Behavior*, Ed. Clifford Madsen and Carol Prickett, pp. 99-114.

Uszler, M. & Larimer, F. (1986). *The piano pedagogy major in the college curriculum*. Princeton, New Jersey: National Conference on Piano Pedagogy.

Walkington, J., Christensen, H.P., & Kock, H. (2001). Developing critical reflection as a part of teaching training and teaching practice. *European Journal of Engineering Technology, 26*(4), 343-350.

Yarborough, C. (1987). The relationship of behavioral self-assessment to the achievement of basic conducting skills. *Journal of Research in Music Education, 35* (3), 183-89.

Yarborough, C., Wapnick, J., and Kelly R. (1979). Effects of videotape feedback techniques on performance, verbalization, and attitude of beginning conductors. *Journal of Research in Music Education, 27,* 103-112.

Yarborough, C. & Hendley, P. (1999). The effect of observation focus on evaluations of choral rehearsal excerpts. *Journal of Research in Music Education 47*(4), 308-318.

Appendix A

Observation #1

1. What aspects of musical performance did you work on?
    - List each topic that you covered in order of presentation (i.e. hand position, phrasing, etc).
    - Indicate whether there was an observable improvement in student performance for each topic area by marking a+(improved) or -(no improvement) next to each topic on your list.
2. Target a teacher behavior and student behavior that you would like to improve or continue to improve
    - Indicate strategies that you will incorporate into your next lesson with the student
    - Indicate if your previously targeted teacher and student behaviors were improved

Observation #2

Teacher or Student Behaviors: talk, play, sing, clap, write, (indicate any other)
Success: student able to respond/play as instructed by teacher .
Unsuccessful: student not able to respond/play as instructed by teacher

Activity:
Total Time:
Teacher Behaviors:

Student Behaviors:

Aspects of Musical Performance worked on:

# of times Student successful
# of times Student not successful:

Activity:
Total Time:
Teacher Behaviors:

Student Behaviors:

Aspects of Musical Performance worked on:

# of times Student successful
# of times Student not successful:

Activity:
Total Time:
Teacher Behaviors:

Student Behaviors:

Aspects of Musical Performance worked on:

# of times Student successful
# of times Student not successful:

Activity:
Total Time:
Teacher Behaviors:

Student Behaviors:

Aspects of Musical Performance worked on:

# of times Student successful
# of times Student not successful:

Observation #3

Use the computerized observation program (SCRIBE) to record the following aspects of your rehearsals/lessons.

Attach printouts of the SUMMARY and TIMELINE to this document.

TOTAL LESSON TIME from the beginning of your first verbal instruction to the end of the last performance _____.

	Frequency	Total Time	% of Lesson Time	Mean Duration	SD	Rate (n/min)
TEACHER TALK	_____	____:____	_____	____:____	_____	_____
TEACHER PERFORM/DEMO	_____	____:____	_____	____:____	_____	_____
STUDENT PERFORMANCE	_____	____:____	_____	____:____	_____	_____
STUDENT TALK	_____	____:____	_____	____:____	_____	_____
(SECOND PASS—TEACHER VERBALIZATIONS)						
INFORMATION (EXPLAIN)	_____					_____
DIRECTIVES (DO IT)	_____					_____
SPECIFIC FEEDBACK	_____					_____
NONSPECIFIC FEEDBACK	_____					_____
OFF-TASK VERBAL	_____					_____
ADDITIONAL OBSERVATIONS (YOUR CHOICE)						
_____	_____	____:____	_____	____:____	_____	_____
_____	_____	____:____	_____	____:____	_____	_____
_____	_____	____:____	_____	____:____	_____	____

Observation #4

For a 12-15 minute segment in which you are actively engaged in improving some aspect of student behavior:

1. List teacher instructions for each task (there may be several tasks within this segment)--what it is the STUDENTS actually DO.

2. With each instruction monitor the following:
- Student successful or unsuccessful trials
- Repeat/no repeat after unsuccessful trial
- Repeat/no repeat after successful trial
- Follow up instructions forward moving (next step of sequence)
- Follow up instructions backward moving (same step of sequence)
- Goals met in this instructional sequence? Why or why not?

3. Were instructions given for student practice?

4. Congruence of actual sequence with your lesson plan & reasons for deviation(s). Did you do what you planned to do? What changes did you make and why?

5. Target a teacher behavior and student behavior that you would like to improve or continue to improve.
- Indicate strategies that you will incorporate into your next lesson with the student.
- Indicate if your previously targeted teacher and student behaviors were improved.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

A Comparison of Testing Formats for Vocal Sight Reading

Michele L. Henry.
Baylor University

Vocal sight-reading is an area of long-standing interest among music education researchers. The ability to "read and notate music" is among the primary goals for music education (CNAEA, 1992) and for choral music educators in particular. Investigators have examined the role of sight-reading in the choral curriculum (Brendell, 1996; Daniels, 1988; Parrish, 1984; Szabo, 1992), pedagogical techniques associated with sight-reading instruction (Demorest, 1996; Henke, 1984; Houlahan & Tacka, 1990a, 1990b; Riggins & Irving, 1988), and materials used in sight-reading instruction (Anderson, 1989; Frances, 1985; Gordon, 1997; Jones, 1979; Kazez, 1992). Researchers have also investigated sight-reading achievement and factors related to sight-reading achievement (Boyle & Lucas, 1990; Daniels, 1986; Demorest & May, 1996; Henry & Demorest, 1994; Killian, 1991; Lucas, 1994).

Throughout the widespread investigations of these numerous researchers, no standard tool for systematic assessment of vocal sight-reading skill was available. By virtue of the fact that no widely accepted or proven assessment tool existed, individual researchers had either to create their own assessment tool specifically for the study at hand or modify existing tools designed for purposes other than vocal sight-reading.

The Vocal Sight-Reading Inventory (VSRI) was created with the intent of providing such a tool. Henry (2001) developed an assessment tool that identified 28 component pitch skills for vocal sight-reading, taking the form of a pitch skill inventory. Pitch skills included ascending and descending conjunct motion, skips and leaps within chordal elements, cadential patterns, modulatory patterns, and chromatic patterns. The subject's score reflected the total number of pitch skills performed correctly. In addition, the score identified the individual skills the subject did or did not perform successfully, making the VSRI useful as a diagnostic tool.

Scoring procedures for the VSRI constituted a further innovation for vocal sight-reading assessment. Scoring in previous studies awarded credit either for each pitch and rhythm sung correctly (Boyle & Lucas, 1990; Demorest, 1998; Demorest & May, 1995; Henry & Demorest, 1994; Killian, 1991) or by measure (Anderson, 1981; Scott, 1996; Watkins & Farnum, 1962). Resulting scores represented subjects' holistic ability to sight read, without providing information as to the content within the test or the types of different sight-reading skills demonstrated by the subject. In addition to providing less specific information, these scoring procedures also took an inordinate amount of time to complete.

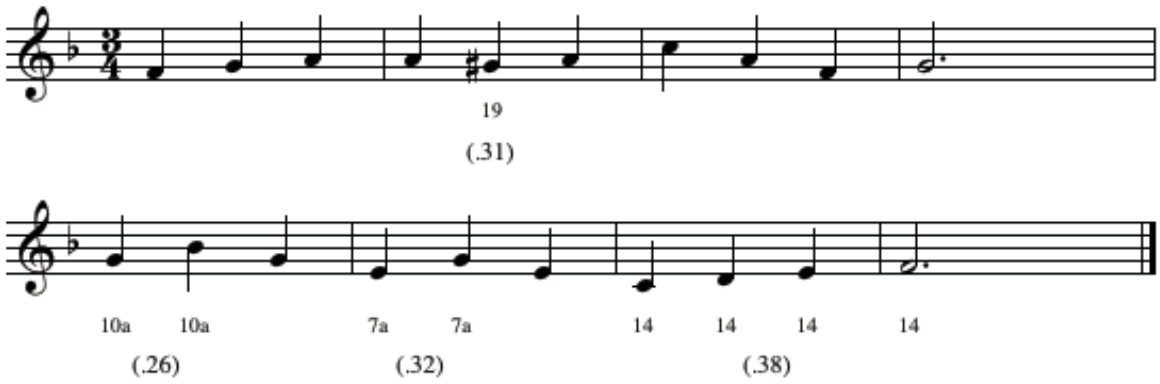
In contrast, VSRI procedures identified one instance of each component pitch skill for scoring. Judges awarded credit for the skill only if every note of the skill was sung correctly. High correlations between VSRI scoring and scoring all pitches in all melodies (r = .96) and high inter-scorer reliability (r = .97) lend credibility to this new scoring system. Additionally, the efficiency of the component skill scoring system, which can be completed in real time, is an important feature for both researchers and classroom teachers.

Henry's study (2001) identified specific component pitch skills within the sight-reading process. Student performance on these skills was measured, and difficulty levels for each skill were obtained. Using the data concerning the difficulty levels of particular vocal sight-reading skills gathered in that study, melodies may now be reconfigured so that skills occur in relative order of difficulty. This is a significant advancement in the ability to test accurately without more difficult skills interfering with the performance of easier skills based on their order in the melody. Melodies can be created that contain skills with a specified range of difficulty or that contain skills in ascending order of difficulty from the easiest to the most difficult within each individual example. These reconfigurations could allow the VSRI to serve many different functions for researchers and classroom teachers.

The Current Study

The purpose of the current study was to compare vocal sight-reading performances on two newly-constructed versions of the VSRI to determine equality of the testing formats. The leveled version contained a total of nine melodies in three sections. Component skills embedded within each melody progressed in difficulty, so that with the exception of cadential skills, skill patterns found in melodies in section one were the easiest, skill patterns found in section two melodies were the next easiest, and skill patterns found in section three melodies were the most difficult, based on the results of previous administrations of the VSRI (Henry, 2001). The comprehensive version contained a total of six melodies. In this version, each melody began with easier skills in the early measures. The skills became progressively more difficult as the melody continued. Cadential skills, regardless of their determined difficulty levels, necessarily appeared at the end of the melodies. Figure 1 contains a melody example from both the leveled and comprehensive testing formats of the VSRI.

Leveled (level two)



Comprehensive

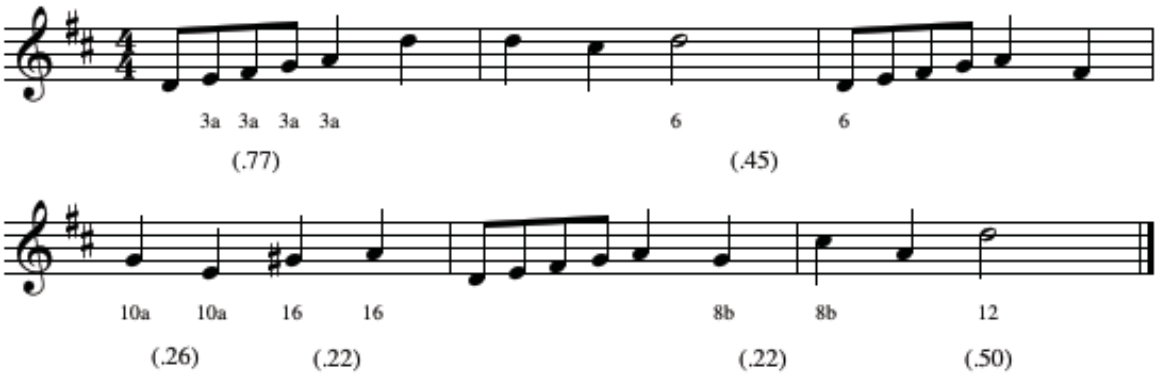


Figure 1. Examples of melodies from the comprehensive and leveled testing formats. Component skill identification numbers appear beneath the pitches. Difficulty levels for each component skill appear in parentheses beneath the component skill number.

Research questions for this study included:

(a) Is there a significant difference in the scores on the leveled and comprehensive testing formats of the VSRI?
(b) What are the difficulty levels for each of the 28 component pitch skills identified on the VSRI? How do these difficulty levels compare with the findings of the original VSRI now that skills have been ordered by difficulty level?
(c) Is there a significant difference for individual pitch skills between testing formats of the VSRI?
(d) Is there a significant difference for individual pitch skills between testing formats of the VSRI?

For the purposes of this study, sight-reading was defined as the accurate singing of pitch in a tonal framework. No effort was made to account for deviations in pitch, such as scooping, sliding, or general flatness or sharpness. Although the melodies appeared in a rhythmic scheme, no assessment of rhythmic accuracy was included in this study.

Method

Study participants were members of choirs in three northeastern Ohio high schools (N = 93). At each location, members from each choir offered for credit were randomly selected to participate in the study. Each of these participants was then randomly administered either the leveled or comprehensive version of the VSRI. The tests were administered by the investigator and an assistant who was an upper-level music education student at a nearby college.

Prior to the test administration, participants completed a questionnaire concerning their musical backgrounds. For testing, individual participants went into a practice room with the investigator or an assistant. Participants were handed a set of melodies and given verbal instructions. For each melody, the key feeling was established through a I-IV-V-I chord progression played on the keyboard in the singer's octave. Participants were given up to 30 seconds to study the melody. When ready, the tape recorder was turned on. The key feeling and tonic note were provided, and the participant sang the melody. At the conclusion of the first melody, the tape recorder was turned off. This procedure was repeated for each melody in the test. The leveled version took approximately

nine minutes to administer, while the comprehensive version took approximately seven minutes to administer.

All scoring occurred from the tape-recorded performances. Scoring occurred in real-time, resulting in an average scoring time of six minutes for the leveled version and four minutes for the comprehensive version. Scoring rules from the VSRI (Henry, 2001) were employed for all subjects. Credit was awarded for a component skill in which each pitch within the skill was performed correctly within a consistent tonal framework. If any pitch within a component skill was sung incorrectly, no credit was awarded for the skill. Both versions of the VSRI contained the same 28 component skills, and therefore 28 points represented a perfect score on each version of the VSRI. Both the investigator and the assistant served as scorers. An inter-scorer reliability analysis revealed a correlation of .98 between scorers. Because of the high correlation between scorers, only the scores of the investigator were used for the purposes of further data analysis.

Results

The mean score for participants completing the comprehensive version of the VSRI was 9.13, with a standard deviation of 6.71. The participants completing the leveled version of the VSRI achieved a mean score of 7.85, with a standard deviation of 5.56. A *t*-test assuming equal variances revealed no significant differences between test versions ($t = 1.00$, $df = 92$, $p < .16$).

The difficulty level for each of the 28 component pitch skills for all participants ($N = 93$) is listed in Table 1. Difficulty levels ranged from .89 (repeated patterns) as the easiest, and .04 (*sol-ti-re-fa* leaps and lowered *mi* implying parallel minor) as the most difficult. A Pearson's *r* analysis revealed a .96 correlation for difficulty levels of individual components skills when compared with the original VSRI (Henry, 2001). Further, a .94 correlation was obtained for component skill difficulty levels between the leveled and comprehensive versions of the VSRI tested in this study.

Table 1

Difficulty Levels for Component Pitch Skills

Conjunct Skills					
Skill		Difficulty	Skill		Difficulty
1	repeated	.89			
2a	*drm* ascending	.78	2b	*drm* descending	.51
3a	*drmfs* ascending	.69	3b	*drmfs* descending	.30
4a	*drmfslt* ascending	.59	4b	*drmfslt* descending	.59
Tonic Skills					
Skill		Difficulty	Skill		Difficulty
5a	*dms* skip	.45	5b	*dms* leap	.45
6	*dd' octave*	.37			
Dominant Skills					
Skill		Difficulty	Skill		Difficulty
7a	*str* skip	.10	7b	*str* leap	.11
8a	*strf* skip	.16	8b	*strf* leap	.04
Subdominant Skills					
Skill		Difficulty	Skill		Difficulty
9a	*fld* skip	.16	9b	*fld* leap	.10
10a	*rfl* skip	.16	10b	*rfl* leap	*.08*
11a	*ldm* skip	.18	11b	*ldm* leap	.11
Cadential Skills					
Skill		Difficulty			
12		end on *d*	.42		
13	*td'*	.23			
14	*sd'*	.30			
15	*sltd'*	.22			
Modulatory Skills					
Skill		Difficulty			
16	V/V secondary dominant	.09			
17	*bm* parallel minor	.04			
Chromatic Skills					
Skill		Difficulty			
18	lower chromatic	.14			
19	upper chromatic	.22			

A X^2-square analysis between testing formats of the VSRI revealed no significant difference in skill difficulty level for 23 of the 28 skills. Significant differences were found for descending patterns *do-re-mi* and *do-re-mi-fa-sol*, skips between *sol-ti-re-fa*, lowered *mi* implying parallel

minor, and upper chromatic patterns. In each case, participants who took the comprehensive version were more successful at completing these tasks.

A multiple regression analysis for the entire sample population ($N = 93$) revealed that the variables of grade, choral experience, keyboard experience, and instrumental experience accounted for 32% of the variance in scores on the VSRI. A significant relationship between performance on the VSRI and both keyboard experience and instrumental experience was established at the .05 level.

Discussion

The overall mean score for both the leveled and comprehensive versions of the VSRI was 8.47 ($N = 93$). The mean scores between testing formats (9.13 for the comprehensive version and 7.85 for the leveled version) did not differ significantly, indicating equivalence between testing forms. This result allows researchers or teachers to select and use whichever version of the VSRI best suits their needs, without concern that the presentation format will affect the results. The comprehensive version of the VSRI is most suitable when the desired result is a complete inventory of the pitch reading ability of the subject. By completing the six melodies contained in the comprehensive version, all 28 component pitch skills are measured. The leveled version of the VSRI is most suitable when the desired result is a score based on a potentially incomplete testing. This scenario might occur when a complete inventory has already occurred and the purpose of testing is to determine achievement since the last testing time. Another plausible reason for conducting an incomplete testing might be for efficiency purposes. With the leveled version of the VSRI, once a subject is unable to successfully complete a section, or makes a pre-determined number of errors within the section, testing ceases. This procedure is similar to that of the *Watkins-Farnum Performance Scale* (1962) for wind instruments, as well as Scott's test of vocal sight-singing for sopranos (1996).

The confidence level with which these testing format options can be undertaken is increased through the additional information gained in this study, particularly in regard to difficulty levels for the individual component skills. The .96 correlation between the difficulty levels on the current versions of the VSRI and the original version indicate a strong stability for the established difficulty levels of each component pitch skill. These results, when considered along with the .94 correlation between pitch skills on the leveled and comprehensive versions of the VSRI, indicate strong validity of the component pitch skills as stable, individual skills that retain the same functional qualities regardless of context.

With the established difficulty levels provided through testing of the initial and revised versions of the VSRI, additional melodies can be constructed and used in alternate versions of the VSRI, provided that component skills occur in proper order. In the case of comprehensive versions, each pitch skill housed within a given melody should increase, rather than decrease, in difficulty through the end of the melody. For leveled versions, pitch skills within a particular melody or section of the test should have generally the same difficulty level, and melodies should be placed in ascending order of difficulty level. Further, pitch skills within existing melodies or exercises can be identified using pitch skill labels from the VSRI. This enables instructors to analyze melodic material based on its pitch skill content, allowing intentional selection and sequencing of melodies or sight-reading examples for use in sight-reading instruction. There are currently several sight-singing textbooks that structure their materials by tonal function of the materials. In particular, the Folk Song Sight Singing Series (Crowe, Lawton, & Whittaker, 1961) designates within each of its ten books the specific tonal and rhythmic tasks contained in each.

In analyzing the discrepancies between the leveled and comprehensive versions of the VSRI for the five component pitch skills that obtained significant differences, the importance of the difficulty level of the skill immediately preceding became apparent. While the progression of the skills in difficulty level throughout the melody was consistent, the difference in the difficulty level of the preceding skills may have interfered with the correlation between forms. For example, skill 2b—*do-re-mi* descending—obtained a .62 difficulty rating on the comprehensive version of the VSRI.

The skill immediately preceding, repeated patterns, obtained a difficulty rating of .87. In the leveled version, skill 2b obtained a .40 difficulty rating, but was preceded by skill 5a, with a .46 difficulty rating. In this case, the difficulty rating obtained for 2b on the comprehensive version was higher than even the skill *preceding* 2b on the leveled version, making it unlikely that subjects would score better on the subsequent skill. This situation occurred for four of the five skills with a significantly different X^2 result (skills 2b, 3b, 8a, and 16). Future versions of the leveled or comprehensive VSRI should attend to these sequential issues to avoid confounding results like the ones obtained in this study. The other skill receiving a significantly different result between forms was skill 17, lowered *mi*, implying the parallel minor. This discrepancy may be accounted for in the difference in presentation between forms. In the comprehensive version, the skill appears in stepwise motion in a manner similar to an upper neighbor chromatic tone. In the leveled version, the skill appears not in stepwise motion, but as an augmented second, which immediately changes direction on the subsequent pitch. The fact that none of the 49 subjects who took the leveled version of the VSRI completed this skill successfully indicates an extremely high level of difficulty for this skill. This result may indicate further subcategories of adjacent and non-adjacent chromatic skills. This possibility should be explored in future versions of the VSRI.

The relationship between keyboard experience, instrumental experience, and success on the VSRI is consistent with results from the original VSRI testing, as well as previous research (Demorest and May, 1996; Henry and Demorest, 1994; Tucker, 1969). While the initial testing of the VSRI found grade and choral experience to have significant relationships with success on the VSRI (Henry, 2001), no such relationship was found in the current study. The strength or lack of strength in these middle school feeder programs could account for this inconsistency. Only one of the three schools in which testing occurred had a feeder program for the high school choirs. Therefore, grade and choral experience may play a less significant role in the results from these participants. Future research in this area should include the use of longitudinal testing of choral students from middle school through high school graduation using the VSRI as an assessment tool. In addition to charting the growth of students' vocal sight-reading achievement, a clearer picture of the role of choral experience and grade level may be obtained.

The comprehensive and leveled testing formats used in the revised versions of the Vocal Sight-Reading Inventory provide choral music educators and researchers with a choice of assessment tools that can be selected based on the particular needs of the user, with the knowledge that both the testing and scoring procedures are consistent, efficient, and informative. While assessing students individually will continue to be a time-consuming proposition, the development in scoring procedures and flexibility of materials and formatting provided by the revised versions of the tool offer significant advances. The ability to score performances in real-time allows for substantially more efficient test administration procedures. The consistency of content and scores between these versions and with previous formats of the VSRI provides confidence to the user of both the validity and reliability of the testing formats. The identification of success for specific component skills provides important diagnostic information, as well as a more complete picture of a subject's specific pitch-reading abilities than any previous test construction format. It is this researcher's hope that future investigations of the various facets of the sight-reading process—whether dealing with student achievement, instructional methodology, materials, or other inquiries—can be conducted with the assistance of the VSRI.

References

Anderson, J. (1981). Effects of tape-recorded aural models on sight-reading and performance skills. *Journal of Research in Music Education, 29*, 23-30.

Boyle, J., & Lucas, K. (1990). The effect of context on sightsinging. *Bulletin of the Council for Research in Music Education, 106*, 1-9.

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*, 1996, 6-14.

Consortium of National Arts Education Associations (1994). *National standards for arts education: What every young American should know and be able to do in the arts*. Reston, VA: Music Educators National Conference.

Crowe, E, Lawton, A, & Whittaker, W. (Eds.). (1961). *The folk song sight singing series*. New York: Oxford University Press.

Daniels, R. (1988). Sight-reading instruction in the choral rehearsal. *UPDATE: Applications for Research in Music Education, 6*, 22-24.

Daniels, R. (1986). Relationships among selected factors and the sight-reading ability of high school mixed choirs. *Journal of Research in Music Education, 34*, 279-289.

Davidson, L., & Scripp, L. (1988). A developmental view of sightsinging. *Journal of Music Theory Pedagogy, 2*, 10-23.

Davidson, L., Scripp, L., & Meyaard, J. (1988). Sightsinging ability: A quantitative and qualitative point of view. *Journal of Music Theory Pedagogy, 2*, 51-68.

Demorest, S. (1998). Improving sight-singing performance in the choral ensemble: The effect of individual testing. *Journal of Research in Music Education, 46*, 182-192.

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

Frances, R. (1985). Tonal principles as teaching principles in music. *Music Perception, 2*, 389-396.

Gordon, Edwin (1997). *Learning sequences in music*. Chicago: GIA Publications.

Henke, H. (1984). The application of Emile Jaques-Dalzroze's solfège-rhythmique to the choral rehearsal. *Choral Journal, 25*, 11-14.

Henry, M. (2001). The development of a vocal sight-reading inventory. *Bulletin of the Council for Research in Music Education, no. 150*, 21-35.

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

Houlahan, M., & Tacka, P. (1990). A suggested sequence for teaching musical elements based on the philosophy of Zoltan Kodály for a college music theory course. *Journal of Music Theory Pedagogy, 4,* 85-109.

Jones, M. (1985). An investigation of the difficulty levels of selected tonal patterns as perceived aurally and performed vocally by high school students. *Dissertation Abstracts International, 40*, 2532A.

Kazez, D. (1992). Solfege drills. *Journal of Music Theory Pedagogy, 6,* 19-34.

Killian, J. (1991). The relationship between sightsinging accuracy and error detection in junior high singers. *Journal of Research in Music Education, 39,* 216-224.

Lucas, K. (1994). Contextual condition and sightsinging achievement of middle school choral students. *Journal of Research in Music Education, 42*, 203-216.

Parrish, D. (1984). *A study of sight-reading pedagogies of six successful high school choral directors in Texas*. Unpublished master's thesis, Baylor University, Waco, TX.

Riggins, H., & Irving, H. (1988). Creative sightsinging. *Journal of Music Theory Pedagogy, 2,* 85-99.

Scott, T. (1996). *The construction of a holistic, criterion-referenced sightsinging test for high school sopranos based on the voluntary national standards for music education.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Szabo, C. (1992). A profile of ten high school choral directors and their activities during one week. (Doctoral dissertation, Kent State University, 1992). *Dissertation Abstracts International, 53*, 08A.

Tucker, D. (1969). Factors related to musical reading ability of senior high school students participating in choral groups. (Doctoral dissertation, University of California) *Dissertations Abstracts International, 31*, 2427A.

Watkins, J., & Farnum, S. (1962). *The Watkins-Farnum performance scale*. Winona, MN: Hal Leonard Music, Inc.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Testing the Predictive Value of Musical Aptitude Scores Upon the Musical Achievement Scores of Beginning Instrumental Students

J. Richard Holsomback, Jr.
Pine Tree Independent School District
Longview, Texas

One of the most important aspects of the instrumental music program is the recruitment of potential students. Instrumental music educators are confronted with the issues of matching the physical characteristics of the potential student with the instrument that they will be best suited within the program. They must also consider the financial ramifications of instrument assignments to potential students. Similarly, these music educators are confronted with maintaining a balanced instrumentation for the good of the overall program.

Before any of these issues are pondered, instrumental music teachers must decide which benchmarks to use in their selection process. This aspect of the recruitment process may be the most important issue in the actual selection process. Should prior musical training in the students' background be considered in the selection process? Should music educators use administrative (principals or counselors) recommendations, or recommendations from their academic teachers? Should an IQ assessment, or other norm-referenced or criterion-referenced academic testing data, such as reading and math scores be considered prior to the selection process? Although music educators have a variety of benchmarks available to them for placement into music education programs, the musical aptitude survey is often the very first benchmark obtained by the educator via the general music instructor.

The testing of musical ability or potential musical achievement dates back to at least 1879 when Wilhelm Wundt opened his institute at the University of Leipzig. His research centered upon testing the creativity of children by testing the reaction time of children to visual, auditory and other sensory stimuli. In the United States musical aptitude testing can be traced to the pioneering work of Carl Seashore. After some 20 years of intensive research, he published the *Measures of Musical Talent* in 1919. Seashore researched musical aptitude issues that included the nature/nurture aspects of musical aptitude, and specific versus global measurements of musical aptitude. In addition, Seashore researched validity of the measurement tools used in assessing musical aptitude and the relationships of the various constructs of musical aptitude to other educational experiences.

Since the Seashore tests were developed, a variety of musical aptitude tests were introduced in the United States. Researchers began to study the constructs of musical aptitude and began to apply their findings to real classroom settings. One of the early pioneers in the constructs of musical aptitude was Rainbow (1963) who worked upon different facets (constructs) of musical aptitude, such as pitch discrimination, rhythmic sensitivity and memory for musical tones. Rainbow concluded that variables outside of music such as academic achievement, academic intelligence and socioeconomic background were significant predictors of musical ability. He also concluded that pitch, rhythm and tonal memory were facets of musical aptitude.

Since the work of Rainbow, other theories of musical aptitude have developed. For example, Karma (1982, 1986) viewed musical aptitude as perceptual and cognitive processing and he questioned the traditional approaches to assessing musical aptitude. He maintained that reliance on simple discrimination tasks as measures of musical aptitude were inappropriate. Karma viewed musical aptitude as a perceptual structuring of music and devised information-processing strategies to assess musical aptitude in terms of cognitive structuring. Gordon (1984, 1987) on the other hand, takes a more developmental approach to musical aptitude. He contends that musical aptitude fluctuates until about the age of nine, after which it becomes more stabilized. The practical problem for the practitioner in the field may be finding a reliable assessment for music aptitude and an assessment that is readily available for classroom use. Most of the older musical aptitude assessments are out of print or very difficult to locate. Similarly, some assessments that are readily available such as *Selmer Music Guidance Survey* (Selmer, Inc., 1980) or the *King Musical Instrument Appraisal* (King, Inc., 1976) have little or no research to support their use in obtaining a musical aptitude benchmark score.

Most of the early investigations linking musical aptitude to musical achievement were actually studies that linked aptitude with academic achievement as measured by standardized tests given to students during their regular academic testing program and not teacher-made assessments as examined in this study. For example, Young (1971) conducted a study to examine the combined powers of three types of standardized test batteries to predict success in an elementary school instrumental music program. The results indicated that Gordon's *Musical Aptitude Profile* and the *Iowa Test of Basic Skills* were the best predictors of success in an elementary school instrumental program when used in tandem.

More recent research has placed more emphasis upon the link between music instructional techniques as measured by standardized music teacher behavior instruments and academic achievement of students using standardized academic tests or musical aptitude assessments to predict potential success in music programs. The results of such research are often not very conclusive (Daone, et. al., 1990; Dryden, 1992; and Corral, 1998). Other more recent research has focused upon the "nonmusical" benefits of participating in musical activities for students (i.e., development of self-discipline, pride in accomplishment, and leadership development). Results of such studies are promising and appear to be more conclusive. Students that participate in school music programs seem to gain various "nonmusical" benefits (Straub, 1994 and Morrison, 1994). Zdzinski (1992) examined student musical aptitude, musical achievement and the "non-school" factors of parental involvement and gender. Zdzinski concluded that there was little correlation between parental involvement and performance achievement, musical achievement or musical aptitude. He did, however, identify a significant relationship for performance achievement among gender, parental involvement and music aptitude.

Another area of research has emerged in the area of music achievement is increasing student musical achievement through the use of various motivational and instructional techniques. The recent research of Vega (2001) developed an action research project to implement a program to improve motivation through the use of multiple intelligences, authentic assessment, technology, and positive teacher feedback to increase levels of student musical achievement. Post-intervention data suggested that all four strategies implemented were effective in increasing student musical achievement. These results are very promising. Further research is certainly warranted in this newer methodology of studying student musical achievement and alternative assessment instruments in the classroom.

Another study using action research methodology was the research of Caliendo and Kopacz (1999). Like Vega (2001) the researchers were investigating student motivation and its link to musical achievement. Their assessments, however, were different. Their review of the literature suggested that students listen and analyze the music to be performed. Student led sectional rehearsals and students critiquing their own videotaped performances in classroom practice were also encouraged. Results indicated that musical achievement and performance scores for band and choral students on the three assessment techniques examined in the action study rose significantly. In sum, early investigations linked musical aptitude with academic achievement as measured by

standardized tests and not teacher-made assessments as examined in this study. Similarly, more recent research has placed emphasis upon the link between music instructional techniques as measured by standardized teacher appraisal instruments and academic achievement of students. Musical aptitude assessments have likewise been employed to predict potential success of students in music programs. The *Selmer Music Guidance Survey* employed in this study, however, is not found in the literature.

The results of linking music instructional techniques as measured by standardized teacher appraisal instruments and academic achievement of students have not led to conclusive implications. Moreover, in the later part of the decade of the 1990s, action-research methodology emerged in the area of music achievement research. The two studies cited were designed to improve student musical achievement/performance levels through the use of various motivational/instructional techniques.

Method

The purpose of this study was to examine the reliability of the *Selmer Music Guidance Survey*. In addition, the study sought to determine the predictive value of those composite musical aptitude scores upon musical achievement, as measured by grades those students actually obtained in their band classes across one academic year. Musical aptitude was defined as scores on the composite score of the *Selmer Music Guidance Survey.* Musical achievement in this study was defined as the grade students received in band classes across six, six-week grading periods, two semester grades and the final average. Their band director took these grades as a part of their regular instrumental music curriculum based on mastery of the state-mandated Texas Essential Knowledge and Skills (TEKS) objectives. A minimum of three assessments was given to each student in the band classes. In some of the six week grading periods there were as many as five assessments taken. In the fifth six-week period, a pass/fail assessment was used because the band was attending a band contest. All assessments were derived from the students' band method books used in their band classes.

The following questions were addressed descriptively and analytically: Were the *Selmer Music Guidance Survey* (1980) composite scores equally reliable as a readily available and tested assessment such as the Gordon (1986) *Intermediate Measures of Music Audiation*? Did the musical aptitude composite scores have predictive value upon the musical achievement scores as measured by grades the sample actually attained in the band classes across one academic year?

The hypotheses stated that the Selmer assessment tool was not as reliable as the Gordon test and there was no difference in the predictive value of the composite musical aptitude scores upon the musical achievement scores of the sample employed in the study. For this study, it was assumed that comparable data were available on the number of pupils in the district and that data on the variables, as defined in this study, were available for all students in a standardized format. Likewise, this study was limited to one East Texas Independent School District, one academic year of musical achievement. Similarly, definitions of the variables that defined musical achievement and musical aptitude were important because other variables and school site data could yield very different results. The basic design was a correlation (predictive) study used to determine the predictive value between musical aptitude composite scores and musical achievement of beginner band students.

The sampling technique was a purposive technique. The intact sample consisted of 91 (n = 91) sixth grade band students in an East Texas school district. According to Gay (1981) 30 subjects are considered to be a minimally acceptable sample size for a correlation study. This sample met and exceeded that particular benchmark in the literature. Students were given the *Selmer Music Guidance Survey* to serve as a guide to instrument assignment and general assessment of the individual needs of the selected band students at the beginning of the school year. Students were then interviewed for physical characteristics to match those characteristics for instrument placement in the sixth grade year, the year in which these students began their instrumental music training.

Since no research was available in the literature for the *Selmer Music Guidance Survey*, the reliability coefficient was computed on a split-half procedure and the Spearman-Brown prophecy formula was used for correction purposes.

The Pearson Product Moment correlation coefficient was computed, examining the relationship between the *Selmer Music Guidance Survey* composite scores and the musical achievement scores of the band students in the purposive sample. The significance level was set at .01 ($p < .01$). If the correlation coefficient met or exceeded the critical value and the relationship was positive, a simple linear regression model was built to test the predictive value of the independent variable (musical aptitude scores) upon the dependent variables (grades received by the band students throughout one academic year). Using the VasserStats (2002) package, the correlation coefficient, the coefficient of determination, slope, y intercept, standard error and the *p* values were obtained to test the predictive value of the simple linear regression.

Results

The reliability coefficient of the *Selmer Music Guidance Survey* composite scores was .72 ($n = 336$) set on an entire sample of a fifth-grade class during the 2000-2001 school year. With the Spearman-Brown formula, the coefficient was .84 on a split-half reliability coefficient technique. The coefficient compared favorably with the Gordon (1986) *Intermediate Measures of Music Audiation* composite score, although the Gordon sample was larger. The Gordon survey reliability coefficient was .80 on a split-half technique ($n = 752$).

The Pearson Product-Moment correlation technique was computed, examining the relationship between the *Selmer Music Guidance Survey* (1980) composite scores and the musical achievement scores of the band students in the purposive sample. The significance level was set at the .01 level ($p < .01$). If the correlation coefficient met or exceeded the critical value and the relationship was positive, a simple linear regression was utilized to establish the predictive ability of the *Selmer Music Guidance Survey* (1980) upon the achievement scores of the beginner band students.

The correlation coefficients for musical aptitude and musical achievement as defined in the study ranged from .290 to .589. All correlations were positive and significant at the .01 level for all six six-week marking periods, the two semester grades and their final grade in the band class. The correlations for the Selmer and six-weeks one, two and three were .340, .589, and .348 respectively. Likewise, the correlations for the Selmer composite scores and the semester average were .539. The correlations for the Selmer and six-weeks four, five and six were .543, .290, and .378 respectively. Similarly, the correlations for semester two were .478 and the correlation for the final averages was .529. These results are shown in Table 1.

Table 1

Product Moment Correlation Coefficients for the Selmer Music Guidance Survey Composite Scores and students' band grades for each grading period across one academic year

Grading Period	Selmer Composite Scores
Six Weeks 1	.340**
Six Weeks 2	.589**
Six Weeks 3	.348**
Semester 1	.539**
Six Weeks 4	.543**
Six Weeks 5	.290**
Six Weeks 6	.378**
Semester 2	.478**
Final Grades	.529**

Note. Source: *Selmer Music Guidance Survey Composite* Scores, Student Grade Reports across one academic year.
** $p < .01$

A linear regression was then employed to test the predictive ability of musical aptitude, as defined by composite scores on the *Selmer Music Guidance Survey* (1980), to the grades that students received in the actual band classes across one academic year. For the Selmer composite scores and the first six weeks, the R^2 was .1161, the slope .116, the Y intercept 80.289 and the standard error was 4.750. For the second six weeks the R^2 was .3469, the slope .397, the Y intercept 59.581 and the standard error was 5.6571. For the third six weeks the R^2 was .1215, the slope .165, the Y intercept 81.604 and the standard error was 4.627. Likewise, for the semester the R^2 was .2915, the slope .242, the Y intercept 73.8113 and the standard error was 3.9267. All regressions were positive and significant at the .01 level for the first semester. The lowest standard error of estimate was for the predictive models during the first semester was found in the semester grades and the Selmer composite scores (3.9267). The linear regression models continued for the rest of the school year.

For the Selmer composite scores and the fourth six weeks, the R^2 was .2957, the slope .293, the Y intercept 68.9872 and the standard error was 4.6966. For the fifth six weeks the R^2 was .0842, the slope .150, the Y intercept 86.5622 and the standard error was 5.1472. Likewise, or the sixth six weeks, the R^2 was .1390, the slope .268, the Y intercept 70.4567 and the standard error was 6.9216. Similarly, for the semester two grades and the Selmer composite scores the R^2 was .2290, the slope .259, the Y intercept 73.3759 and the standard error was 4.9338. For the final grades the R^2 was .2798, the slope .247, the Y intercept 74.0732 and the standard error was 4.1225. All regressions were positive and significant at the .01 level for the second semester. The lowest standard error of estimate was for the predictive models during the second semester was found in the final grades and the Selmer composite scores (4.1225). The results for the regression models are shown in Table 2.

Table 2

Simple Linear Regression models for the Selmer Music Guidance Survey Composite Scores and grades in band classes across one academic year

	R	R^2	Slope	Y Intercept	St Error of Estimate	*P* value**
Six Weeks 1	.340	.116	.166	80.2892	4.7560	.00094
Six Weeks 2	.589	.346	.3975	59.5811	5.6571	.00010
Six Weeks 3	.348	.121	.1659	81.6049	4.6271	.00070
Semester 1	**.539**	**.291**	**.2428**	**73.8113**	**3.9267**	**.00010**
Six Weeks 4	.543	.295	.2934	68.9872	4.6966	.00010
Six Weeks 5	.290	.084	.1504	86.5622	5.1472	.00520
Six Weeks 6	.378	.139	.2681	70.4597	6.9216	.00020
Semester 2	.478	.229	.2592	73.3759	4.9338	.00010
Final Grades in Band	.529	.279	.2478	74.0732	4.1225	.00100

Note. Source*: Selmer Music Guidance Survey* and grades students received in band across one academic year
** All *P* values met or exceeded the .01 level ($p < .01$)
Those in bold type indicate standard error estimates values that are the lowest in all of the regression models.

Discussion

The first analysis that was performed on the data in this study was a correlation methodology to determine the reliability coefficient of the *Selmer Music Guidance Survey* employed in this study. The split-half method was used to establish the reliability coefficient and Spearman-Brown prophecy formula was also used for correction purposes. Likewise, the standard error of measure was computed for the *Selmer Music Guidance Survey*, since no literature existed for that musical aptitude test. The findings of the reliability coefficient analysis indicated that: 1. The reliability coefficient of the *Selmer Music Guidance Survey* composite scores was .84 ($n = 336$) set on an entire sample of a fifth-grade class during the 2000-2001 school year and that coefficient compared favorably with the Gordon (1986) *Intermediate Measures of Music Audiation* composite score, although the Gordon sample was larger. The Gordon survey reliability coefficient was .80 on a split-half technique ($n = 752$).

Based on the findings in the reliability coefficient analysis, the research question asked, "Were the *Selmer Music Guidance Survey* (1980) composite scores equally reliable as a readily available and tested assessment such as the Gordon (1986) *Intermediate Measures of Music Audiation*?" It

can be stated that the Selmer composite scores compare favorably with the Gordon composite scores test, keeping in mind that the Gordon test had a larger number in its sample size. The hypothesis stated that "the Selmer was not as reliable as the Gordon test'. The Selmer test compares favorably with the Gordon as long as the researcher keeps in mind that the sample sizes were different and further research with the Selmer composite need to be undertaken to see if the coefficient holds true under a larger sample size.

The second analysis that was performed on the data in this study was a simple linear regression methodology to determine if the *Selmer Music Guidance Survey* employed in this study had predictive ability upon music achievement as defined as grades students received in their band classes across one academic year. The findings of the simple linear regression analysis indicated that:

> All regressions were positive and significant at the .01 level for the first semester. The lowest standard error of estimate was for the predictive models during the first semester was found in the semester I grades and the Selmer composite scores (3.9267).
>
> All regressions were positive and significant at the .01 level for the second semester. The lowest standard error of estimate for the predictive models during the second semester was found in the final grades and the Selmer composite scores (4.1225).

Based on the findings in the linear regression analysis, the research question asked, "Did the musical aptitude composite scores have predictive value upon the musical achievement scores as measured by grades the sample actually attained in the band classes across one academic year?" It can be stated that the Selmer composite scores did have predictive value upon the grades students received in their band classes across one academic year.

Implications

Since the Selmer instrument does compare favorably to the Gordon, more research is needed to determine if a larger sample size affects the reliability coefficient. Since no prior research has been conducted on the Selmer, the results of this study indicated that it is reliable to use as a benchmark.

Although no casual relationships can be determined from the predictive analysis, the Selmer composite scores have positive and significant relationships with the students' band grades. These relationships appear to have predictive value, as well. Use of the *Selmer Music Guidance Survey* as a benchmark for success, based on the findings of this study appears to be one predictor of musical achievement.

References

Caliendo, E. and Kpoacz, R. (1999). *Improving student motivation and performance in music programs* (Report No. S0031225). Illinois: Saint Xavier University, Master's Research Project, and (ERIC Document Reproduction Service No. ED 437 316).

Corral, S.J. (1998). *A comparison of the California Tests of Basic Skills between fourth and fifth grade instrumental music pullout students and students not involved in the instrumental music program* (Report No. TM029726). US: Maryland. (ERIC Document Reproduction Service No. ED 430 013).

Doane, C., et. al. (1990). A validation of music teacher behaviors based on music achievement in elementary general music students, *Research Perspectives in Music Education, 44* (1), 24-41.

Dryden, S. (1992). *The impact of instrumental music instruction on the academic achievement of fifth grade students*, (Report No. S0023799). Fort Hays, KS: Fort Hays State University Master's Thesis. (ERIC Document Reproduction No. ED 368 634).

Gay, L.R. (1981). *Educational research: Compentencies for analysis and application.* Columbus, OH: Merrill.

Gordon, E. (1984). *Learning sequences in music*, Chicago, IL; GIA Publications.

Gordon, E. (1986). *Intermediate measures of music audiation*, Chicago, IL: GIA Publications.

Gordon, E. (1987). *The nature, description, measurement and evaluation of music aptitude*, Chicago, IL: GIA Publications.

Karma, K. (1982). Validating tests of musical aptitude, *Psychology of Music, 10* (1), 33-36.

Karma, K. (1986). Item difficulty values in measuring components of musical aptitude, *Bulletin of the Council of Research in Music Education, 89*, 1-13.

King Corporation (1976). *King Musical Instrument Appraisal Test*. Eastlake, OH: King Corporation.

Morrison, S. (1994). Music students and academic growth, *Music Educators Journal, 81*(2), 33-36.

Rainbow, E.L. (1965). *A pilot student to investigate the constructs of musical aptitude*, Unpublished doctoral dissertation, State University of Iowa.

Selmer Corporation (1980). *The Selmer Music Guidance Survey*, Elkhart, IN: The Selmer Corporation.

Straub, D. (1994). Music as an academic discipline: Breaking new ground, *NASSP Bulletin, 78* (561), 30-33.

Young, W. (1971). The role of musical aptitude, intelligence, and academic achievement in predicting the musical attainment of elementary instrumental musical students, *Journal of Research in Music Education, 19* (4), 385-398.

VasserStats (2002). *VasserStats: A web-based statistical package for the behavioral sciences*, Richard Lowry, author, Poughkeepsie, NY: Vasser College. http://faculty.vassar.edu/lowry/VassarStats.html.

Vega, L.(2001). *Increasing student music achievement through the use of motivational strategies*, (Report No. S0022804). Illinois: Saint Xavier University, Master of Arts Action Research Project, Field-Based Masters Program, (ERIC Document Reproduction Service No. ED 456086).

Zdzinski, S. (1992). Relationships among parental involvement, music aptitude, and musical achievement of instrumental music students, *Journal of Research in Music Education, 40* (2), 114-125.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Perceived Complexity, Likeability, and Familiarity Ratings in Four Chopin Piano Pieces between Music Majors and Non-Music Majors

Allison Hudak
Princeton, New Jersey

Many studies in the psychology of aesthetics have focused on the relationship between the novelty and complexity of a stimuli and its hedonic liking of the listener. Several conflicting theories have been proposed (Cantor, 1968; Zajonc, 1968), but probably the most widely-accepted is that of Berlyne (1970, 1971). Berlyne (1970, 1971) proposed that an inverted-U shaped relationship exists between these two variables. Berlyne drew on the Yerkes-Dodson hypothesis (Yerkes and Dodson, 1908) in proposing that aesthetic judgments reflect our attempts to optimize our psychobiological arousal level, which is directly related to hedonic tone, or liking. Artistic stimuli can increase or decrease this arousal level, and the inverted-U theory predicts that we prefer stimuli that give rise to intermediate rather than to very high or very low levels of arousal. Smith and Cuddy (1986) proposed three assumptions about aesthetic preference: (1) The critical aspect of a stimulus that determines its affective value is its complexity, (2) the relationship between complexity and affective value can be described by the inverted-U curve: an intermediate level of complexity elicits the maximum affect while lower and higher levels of complexity elicit less affect, and (3) the effect of stimulus exposure lowers its complexity and alters the affective values of stimulus patterns (repetition, training, or practice).

Current research focuses on the formal variables of musical stimuli like complexity, novelty, familiarity, redundancy, uncertainty, and orderliness. Various measures of aesthetic response include liking, showing interest, and the familiarity and complexity of the stimulus (Hargreaves, 1986). Vitz (1966) found an inverted-U relationship between the information content of tone sequences and subjects' ratings of their pleasantness. Heyduk (1975) found support for his own optimal complexity model that predicted that listeners like music stimuli that are of optimum complexity to them more than those that are either too complex or too simple at any given point in time. This model also predicts that repeated exposure to a stimulus increases its familiarity and should lead to a reduction in the perceived complexity of that stimulus. The model helps predict how a listener's liking for a piece should change as a function of repetition. In summary, the optimal complexity and preference-feedback hypotheses suggest that ratings of the same music excerpts for subjective complexity or familiarity should produce different relationships with liking at any given point in time (North & Hargreaves, 1995).

The relationship for which there is little empirical evidence is between subjective complexity and familiarity. Two possible relationships may exist between these two variables: increasing familiarity by means of repetition should give rise to a decrease in subjective complexity or when considering the possible effects of subjective complexity on familiarity, a reverse causal relationship leads to the hypothesis of an inverted-U relationship between the two variables (North & Hargreaves, 1995). When considering the second proposition, it is according to the optimal complexity model that most listeners should dislike excerpts of low and high subjective complexity, and the preference-feedback mechanism should lead to their avoidance, giving rise to low familiarity ratings in each case. Moderately complex excerpts should be liked, causing

listeners to seek further exposure to them that in turn should give rise to relatively high familiarity ratings. Complexity may be either objective, a function of the stimulus attributes varied according to some systematic and objective procedure, or subjective, the apparent perceived complexity that is assumed to be a function of the interaction between the stimulus's objective complexity and the listener's musical knowledge, experience with the musical style, and familiarity with the particular musical stimulus (Hargreaves, 1986, pp. 116-117). In reality, most collative variables are relativistic, meaning their qualities depend on the interaction between the structural attributes of the musical stimulus and the listener's prior experience with music in general, with the style, and with the particular piece (Radocy and Boyle, 1997).

In a study conducted by North and Hargreaves (1995), subjects rated sixty excerpts of popular music for liking, subjective complexity, or familiarity. Results showed a positive relationship between liking and familiarity and an inverted-U relationship between liking and subjective complexity. McMullen and Arnold (1976) determined whether the same factor had an influence on preference and interest responses for rhythmic sequences. Results suggested the beginning of a curve showing preference to be an inverted-U shaped function of the amount of distributional redundancy in rhythmic sequences and that interest responses appeared to be an increasing monotonic function of distributional redundancy in rhythmic sequences.

Other studies such as those by Munsinger and Kessen (1964) and Day (1967, 1968) have found preference to be a multi-modal function of complexity. These studies have found that as stimulus complexity increases, preference responses will rise to a peak and then decline as complexity continues to increase. Peak levels of preference are usually found at intermediate levels of complexity.

The inverted-U theory is useful to advertisers that use music to stimulate the consumer. By analyzing familiar and unfamiliar melodies, Hahn and Hwang (1999) found that familiarity of music significantly affected the message recall in television advertising. The effect of the tempo of the music was also significant. Tempo affected music-message congruency and attention gaining. The prediction of an inverted-U was fully supported when familiar music was used, but not when unfamiliar music was used in the backgrounds of television commercials.

This theory is also interesting when analyzing popular music. Brentar, Neuendorf, and Armstrong (1994) found an inverted-U relationship between frequency of exposure to rock and popular songs and affective responses to the songs among undergraduate students. The greatest affective responses occurred around eight exposures to the piece, but after eight exposures, a downward trend occurred. This type of information could be useful to radio disc jockeys that repeat certain popular songs continuously. Perhaps they should consider the inverted-U theory when playing a minimal selection of pieces to their audiences and instead, increase their selections so audience members will still find new "hit" songs engaging after repeated exposures.

In a study by Hargreaves (1984), repetition on liking for music supported the inverted-U subjective complexity hypotheses that states liking for a piece of music is at its maximum at an optimum intermediate level of subjective complexity, after subjects rated their preferences for liking popular, classical, and avant-garde jazz music excerpts. Popular music tended to reach the maximum of pleasantness at an early repetition whereas classical selections reached their affective height with later performances. This could support the idea that classical music lasts longer than popular music (i.e., Beethoven's Fifth Symphony has remained on the radio longer than the Goo Goo Doll's "Black Balloon.").

In another study, Burke and Gridley (1990) found an inverted-U curve was obtained when forty college students plotted their personal preferences against complexity for musical excerpts by Bach, Debussy, Grieg, and Boulez. They found that the more complex the excerpt, the lower the likeability scores. Those excerpts with intermediate complexity received the highest likeability scores.

In summary, the optimal complexity and preference-feedback-hypotheses predicts an inverted-U relationship between subjective complexity and liking, provided that a sufficiently wide range of stimulus complexity is represented.

Statement of the Problem

The main contribution of the present study is to assess the relationship between familiarity, subjective complexity, and liking for excerpts of solo piano music composed by Frederic Chopin between music majors and non-music majors.

Method

The present study seeks to discover if an inverted-U relationship exists between four different complexity levels of carefully graded pieces of solo piano literature as graded in Magrath's (1995) *The Pianist's Guide to Standard Teaching and Performance Literature* and its likeability in two sections of group piano (non-music majors and music-majors (non-pianists) from The University of Texas at Austin.

Music excerpts

The choice of four music excerpts for this study was determined by three primary considerations: (1) that they should all be representative of a homogenous style or genre, thus minimizing the possibility of subjects' biases influencing responses; (2) that they should represent a wide range of complexity/familiarity within this style; and (3) that all pieces chosen were in a major key signature by the same composer, thus minimizing the possibility of subjects' biases influencing responses based on tonality or style preferences. All pieces were selected from the Magrath (1995) guide since they had all been carefully graded as standard teaching repertoire for students. Magrath's grading system ranges from one to ten, ten being the most complex either technically, musically, or a combination of both. Three doctoral music education students listened to the excerpts to determine if the grading matched their intuition of the complexity of the examples. The four music excerpts chosen were solo piano literature composed by Frederic Chopin: *Prelude, op. 29, no. 7 in A* (level 6), *Mazurka op. 7, no. 5 in C* (level 7), *Nocturne, op. 9, no. 2 in Eb* (level 9), and *Etude op. 10, no. 5 in Gb* (> level 10). The first thirty seconds of each piece performed by Vladimir Ashkenazy was recorded from a compact disc to an audio-cassette tape. Ten-second intervals of silence were recorded between each excerpt. The order of the pieces was randomly assigned.

Subjects

The subjects were 28 undergraduate students from The University of Texas at Austin enrolled in either group piano for non-music majors or group piano for music majors (non-pianists). Fifteen subjects participated from the non-music-majors group piano section and thirteen subjects participated from the music-majors group piano section.

Design and procedure

All subjects were read the same directions before listening to the music excerpts.

> You will listen to a tape of pre-recorded piano music. You will listen to the tape twice. You will hear thirty seconds of music and then ten seconds of silence. During the moments of silence, mark your sheet. Listen once through the tape to get an idea of the musical selections. After each excerpt, you will have ten seconds to circle one number for item c for each musical excerpt for "familiarity." Listen again through the tape and circle one number for "how much you liked" and "complexity" of each musical excerpt (items a and b). When you are finished, turn in your sheet to me.

Students were allowed to ask questions for clarification of the directions and the specific items being rated. After the questions, the tape was played for the students. Using 11-point Likert-type scales (1 = low, 11 = high), all subjects rated each excerpt for familiarity, liking, and complexity. Each excerpt was rated immediately after its presentation.

Results

The mean ratings of complexity for the four excerpts in the non-music majors group piano were 5.4 (prelude), 9.6 (mazurka), 7.0 (nocturne), and 10.33 (etude). The mean ratings of complexity for the four excerpts in the music-majors group piano section were 4.31 (prelude), 8.23 (mazurka), 6.85 (nocturne), and 9.62 (etude). Both groups rated the *Prelude* as the least complex and the *Etude* as the most complex. Both groups also rated the *Mazurka* more difficult than the *Nocturne.* This could be due to the tempo. The *Mazurka* was performed at a very fast tempo while the *Nocturne* was performed at a much slower tempo with more lyricism. The following chart shows the mean ratings of complexity for the four graded pieces.

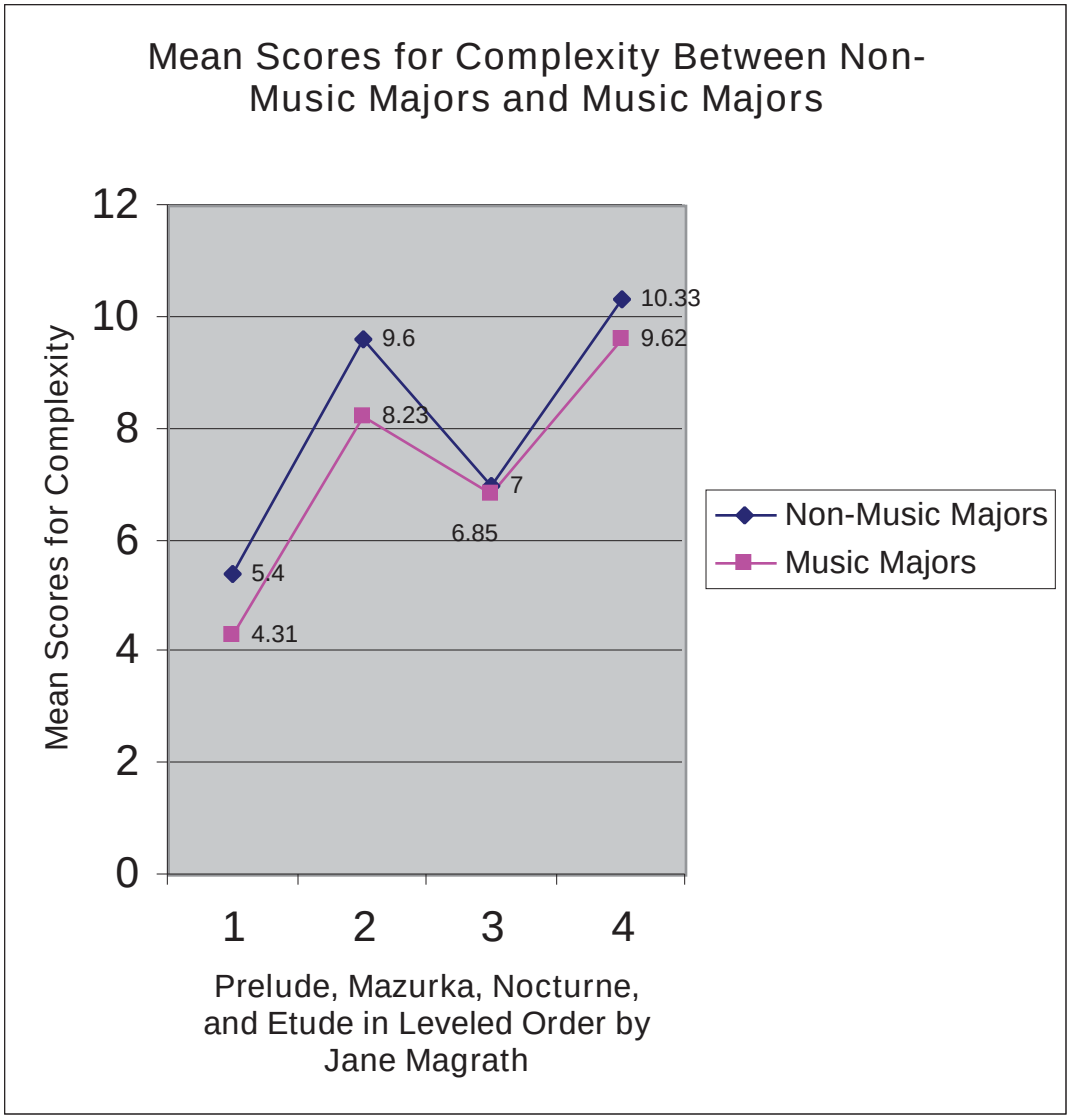


Figure 1. Mean scores for complexity between non-music majors and music majors.

Regarding likeability ratings, the mean ratings for the non-music majors were 7.53 (prelude), 6.86 (mazurka), 9.73 (nocturne), and 9.06 (etude). The mean ratings for the music-majors were 6.85 (prelude), 6.62 (mazurka), 9.46 (nocturne), and 7.92 (etude). Both groups' mean likeability ratings showed that the *Mazurka* was rated the lowest for likeability, and that the *Nocturne* was rated the highest for likeability. When likeability scores were plotted against the perceived complexity determined by Magrath (1995), a slight inverted-U shape was present between the *Mazurka*, the *Nocturne*, and the *Etude* indicating that intermediate complexity yields the greatest likeability ratings. In both instances, the *Nocturne* was rated the highest for likeability and rated an intermediate level of complexity by both groups. The following chart shows the mean ratings for likeability against the leveled pieces as determined by Magrath (1995).

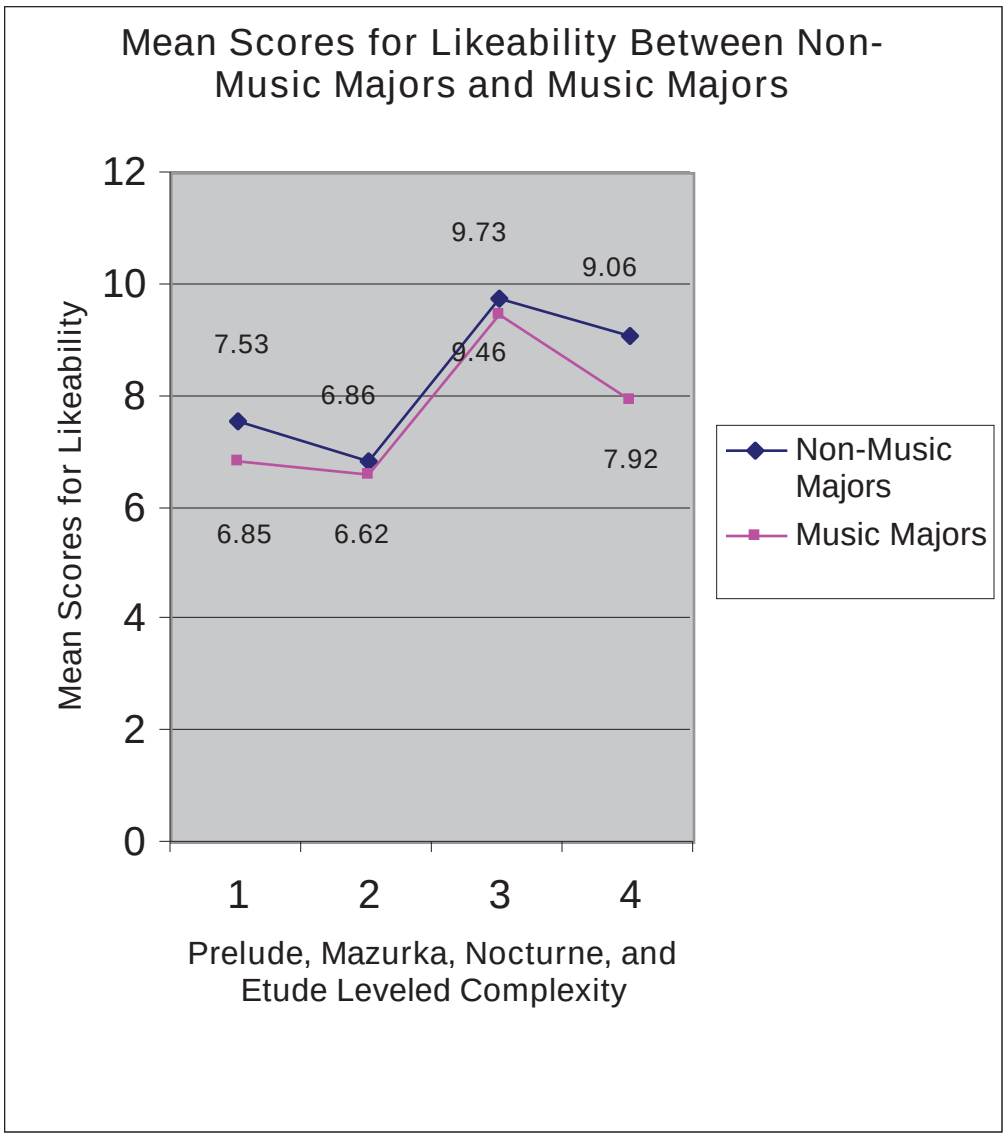


Figure 2. Mean scores for likeability between non-music majors and music majors.

Regarding familiarity, the mean ratings in the non-music majors' section were 5.67 (prelude), 1.93 (mazurka), 7.67 (nocturne), and 2.27 (etude). The mean ratings in the music-majors' section were 5.08 (prelude), 2.23 (mazurka), 6.38 (nocturne), and 2.27 (etude). In this instance, both sections rated the pieces in a similar relationship showing a slight inverted-U shape between the *Mazurka*, the *Nocturne,* and the *Etude* indicating that the most familiar piece of music yielded the highest familiarity ratings. The following chart shows the mean familiarity ratings from the non-music majors' section and the music majors' section of group piano.

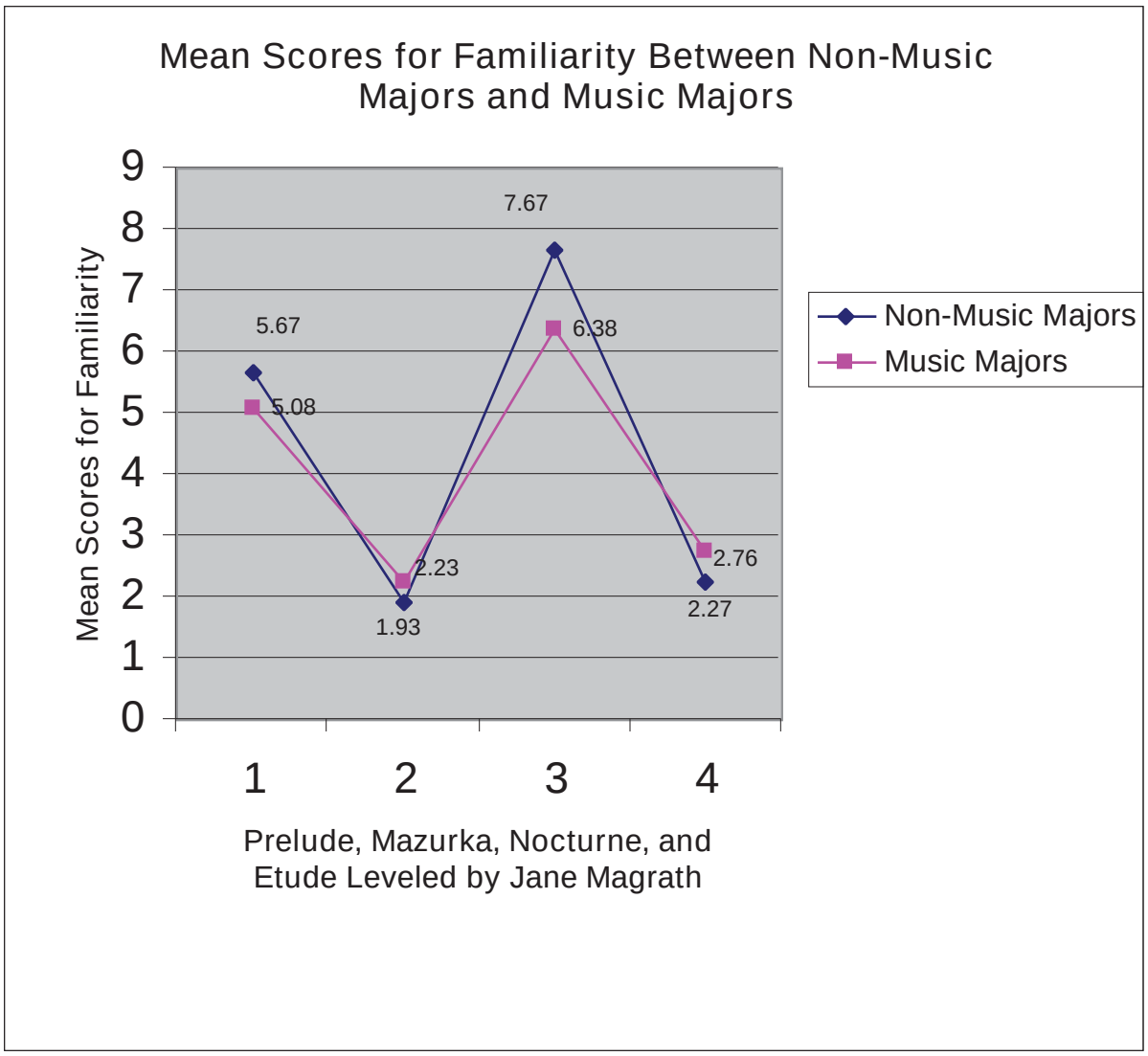


Figure 3. Mean scores for familiarity between non-music majors and music majors.

Discussion

A 2 x 4 Two-Factor ANOVA with Repeated Measures on One Factor Test for likeability, complexity, and familiarity ratings between non-music majors and music majors was conducted. Results indicated that for likeability, there were no significant differences between the non-music majors and the music majors' rankings for each piece. Both classes ranked the pieces very similarly according to their likeability preferences. Analysis of likeability among pieces did support a significant difference between at least one of the pieces and their likeability rankings indicating that one piece of music was preferred more than another in at least one instance where $p< 0.0001$. There were no interaction differences between the two groups of piano students and their rankings for individual pieces of music.

Regarding complexity, results indicated that there were no significant differences between the two groups of students. Both groups of students rated each piece very similarly. Results did show a significant difference in complexity ratings between each individual piece indicating that at least one piece of music was significantly perceived more complex than the other pieces in at least one instance where $p< 0.0001$. There were no significant differences in the interaction between the groups of piano students and their complexity ratings.

For familiarity, results indicated that there was no significant difference between the non-music majors and the music majors group piano sections. Both groups were approximately equal in their familiarity of the individual pieces of music. Results did show that there was a significant difference between at least one piece of music in its familiarity in at least one instance in comparison with the other pieces of music where $p<0.001$. There was not a significant difference in the interaction of the groups and their ratings for familiarity for each individual piece of music.

In summary, there is a slight inverted-U shape for likeability, complexity, and familiarity ratings with both groups. It is interesting to note that although the doctoral reliability observers noted the "correct" complexity ratings as suggested by Magrath (1995) the music majors rated the pieces in the same manner as the non-music majors placing the order of complexity from the prelude to the nocturne to the mazurka to the etude instead of the prelude to the mazurka to the nocturne to the etude. Perhaps the fast tempo used in the mazurka masked the complexity of the piece and students who were unfamiliar with the piece thought that the complexity of the piece was derived from the speed instead of the harmonic and melodic content. Because the nocturne was very slow and lyrical in nature, perhaps students thought this yielded a less complex piece of music.

Both groups of students preferred the nocturne to any of the other pieces. This supports the inverted-U theory in that intermediate complexity yields the greatest likeability ratings. However, this piece was also ranked the most familiar piece of music which could have biased the ratings for likeability.

Further study is needed utilizing the inverted-U and carefully graded and sequenced pieces of music. This study was interesting in that it focused on one composer, one instrument, one genre, one stylistic period, and that the ratings according to Magrath (1995) were only a few leveled ratings from each other. Many studies have focused on using pieces of music that are obviously spread in complexity whereas this study focused on a narrow range of complexity and still yielded similar inverted-U results for three of the pieces.

References

Berlyne, D.E. (1970). Novelty, complexity and hedonic value. *Perception and Psychophysics, 8,* 279-286.

Berlyne, D.E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Brentar, J.E., Neuendorf, K.A., & Armstrong, G.B. (1994). Exposure effects and affective responses to music. *Communication Monographs, 61(2),* 161-181.

Burke, M.J. & Gridley, M.C. (1990). Musical preferences as a function of stimulus complexity and listeners' sophistication. *Perceptual & Motor Skills, 71(2),* 687-690.

Cantor, G.N. (1968). Children's "like-dislike" ratings of familiarized and unfamiliarized visual stimuli. *Journal of Experimental Child Psychology, 6,* 651-657.

Chopin, Frederic. (1995). *Favourite piano works: Beliebte klavierwerke.* Performed by Vladimir Ashkenazy. London: The Decca Record Company Limited.

Day, H. (1967). Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. *Perception and Psychophysics, 2,* 281-286.

Day, H. (1968). The importance of symmetry and complexity in the evaluation of complexity, interest, and pleasingness. *Psychonomic Science, 10(9),* 339-340.

Hahn, M. & Hwang, I. (1999). Effects of tempo and familiarity of background music on message processing in TV advertising: A resource-matching perspective. *Psychology & Marketing, 16(8),* 659-675.

Hargreaves, D.J. (1984). The effects of repetition on liking for music. *Journal of Research in Music Education, 32(1),* 35-47.

Hargreaves, D.J. (1986). *The developmental psychology of music.* Cambridge: Cambridge University Press.

Heyduk, R.G. (1975). Rated preference for music composition as it relates to complexity and exposure frequency. *Perception and Psychophysics, 17,* 84-91.

Magrath, J. (1995). *The pianist's guide to standard teaching and performance literature*. Van Nuys, CA: Alfred Publishing Co., Inc.

McMullen, P.T., & Arnold, M.J. (1976). Preference and interest as functions of distributional redundancy in rhythmic sequences. *Journal of Research in Music Education, 24(1),* 22-31.

Munsinger, Harry & Kessen, W. (1964). Uncertainty, structure and preference. *Psychological Monographs, 78(9),* whole no. 586, 1-24.

North, A.C. & Hargreaves, D.J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology, 14(1),* 77-93.

Radocy, R.E., & Boyle, J.D. (1997). *Psychological foundations of musical behavior*. Springfield, IL: Charles C. Thomas, Publisher, Ltd.

Smith, K.C., & Cuddy, L.L. (1986). The pleasingness of melodic sequences: Contrasting effects of repetition and rule-familiarity. *Psychology of Music, 14,* 17-32.

Vitz, P.C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology, 71,* 74-79.

Yerkes, R.M., & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurological Psychology, 18,* 459-482.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9* (Monograph Supplement 2, Part 2), 1-21.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Multicultural Education: The Influence of Tanglewood Symposium on the Use of Ethnic Music in Selected Fifth, Sixth, and Seventh Grade Silver Burdett Music Series Textbooks

Sibel Karakelle
University of North Texas

America has frequently been described as a cultural "melting pot" because it is made up of a variety of races, cultures, and ethnicities. As more immigrants come to America searching for a better life, the population naturally becomes more diverse. This occurrence has, in turn, sparked a great debate about multiculturalism. When it comes to education, some of the issues under fire include who is benefiting from the education and presenting the material in a way that would offend the least amount of people.

In the 1930s, several educators called for programs of cultural diversity that would encourage ethnic and minority students to study their heritages. This has not been a simple feat because of the great diversity within individual cultures. A look at the 1990 census shows that the American population changed more noticeably in the eighties than in any other time in the twentieth century, with one out of every four Americans identifying himself or herself as Black, Hispanic, Asian, Pacific Islander, or American Indian. The number of foreign-born residents reached an all time high of twenty million, surpassing the 1980 record of fourteen million. [1] Current census studies indicate that by the year 2020 the number of Hispanic descent children under the age of eighteen would increase by 20% and the number of children of African American descent would increase by 15%. By the 1960s, the growing need for multicultural education had been detected. Consequently, teachers and leaders in the field of music education organized the Tanglewood Symposium as a forum in which they could discuss how to react to the multicultural movement and to the needs in music education. Thus, it is relevant to investigate to what degree the content of Silver Burdett Music Series textbooks changed after the Tanglewood Symposium.

Educators and philosophers agree that an important step in successfully joining multiple cultures is to develop an understanding of one another's background.[2] Proponents of

---

1 Gould, K. H. (1995). The misconstruing of multiculturalism: the Stanford debate and Social Work. Social Work, p. 198-204.

2 Ogbu, J. (1995). Understanding cultural diversity and learning. In Handbook of research on multicultural education, edited by J. A. Banks and C. A. McGee Banks. New York: Macmillan.

multicultural education argue that it offers students skills for balanced appreciation and critique of other cultures as well as their own. [3] Yet, one cannot have a true understanding of a subject by only possessing knowledge of one side of it, and there would never be enough time in our current school year to equally cover individual contributions of every nationality. However, by engaging students in activities that follow the guidelines of a multicultural curriculum, educators can open young minds and find ways to make learning exciting as well as educational.

In 1992, the New York State Social Studies Review and Development Committee proposed six guidelines that teachers should follow in order to help break down ethnic barriers. The first proposal was that social studies should be taught from a global perspective. We are all equal owners of the earth; none of us are more entitled to its wealth than others. The uniqueness of each individual is what adds variety to our everyday life. Second, social studies should continue to serve nation-building purposes. By accentuating that which we have in common, it will be easier to examine the elements that make us different. Third, the curriculum must strive to stay informed about the most up-to-date scholarship. With this knowledge, educators can provide students with an excellent support. Fourth, students need to see themselves as active makers and changers of culture and society. If given the skills to judge people and their thoughts fairly and the confidence that they can make a difference, students will demonstrate more responsibility in the future. Fifth, the curriculum should be committed to continuing examination of democratic values as an essential basis for social organization and nation building. Although the democratic system is far from perfect, it has proven in the past that it can be effective if we continue to put effort into maintaining it while leaving it open for change. Finally, social studies should be taught not solely as information but rather through the critical examination of ideas and events rooted in time and place and through responding to social interests. Teachers need to teach with excitement that sparks children's interest and motivates students to want to take part in the shaping of their country's future.[4]

In order to provide a well-rounded multicultural discussion, teachers need to let students know how knowledge reflects the social, political, and economic contexts from which the knowledge has originated (Banks, 1991). The knowledge provided by powerful groups in society differs greatly from that of its less powerful counterparts.[5] Another important aspect students need to realize is that knowledge alone is not enough to shape a society. The members themselves have to be willing to put forth the time and effort and show an interest in shaping their society in order for it to benefit all people.

While generally opposed to the idea of multiculturalism, Francis Ryan (1993)[6] pointed out that "multicultural education programs indeed may be helpful for all students in developing perspective-taking skills and an appreciation for how ethnic and minority traditions have evolved and changed as each came into contact with other groups" (p. 137). It would certainly give people a sense of ethnic pride to know how their forefathers contributed to the building of the American society we enjoy today. It is also a great feeling to know that we can change what we feel is wrong in order to build a better system for our children. Minorities would benefit from

---

3 Stotsky, S. (1992). Academic versus ideological education in the classroom. The Education Digest Mar, 64-66.

4 New York State Social Studies Review and Development Committee (1992). Multicultural education benefits all students: education in America - opposing viewpoints. CA : Greenhaven.

5 Banks, J. (1991). A multicultural literacy and curriculum reform. The Education Digest, 10-13.

6 Ryan, F. J. (1993). The perils of multiculturalism: schooling for the group. Educational Horizons, Spring: 134-38.

learning about the evolution of their culture and from realizing that the ups and downs along the way do not necessarily mean that their particular lifestyle is in danger of extinction.

Some opponents feel that the idea of multiculturalism will, instead of uniting cultures, actually divide them. They feel that Americans should think of themselves as a whole rather than as people from different places all living together. The proponents of this idea go even further to say that multiculturalism actually puts at risk our democratic tradition, the cornerstone of American society.[7] In his article, "Balancing Multicultural and Civic Education Will Take More Than Social Stew,"[8] Paul Gagnon (1991) pointed out that "education in the origins, evolution, advances, and defeats of democracy must, by its nature, be heavily Western and also demand great attention to political history" (p. 8). Since both modern democracy and its alternatives are derived mostly from European past and since most of the educators have traditionally been white males, the educational choices are certainly limited. If we try to avoid these truths or sidestep them in any way, we cannot honestly say we are giving an accurate description of our history.

The importance of multicultural education in music was first addressed at the Tanglewood Symposium, which was held in Massachusetts in 1967. Music Education National Conference (MENC) sponsored the symposium, and its purpose was to discuss the problems of music education related to the cultural, economic, and social changes in the society.[9] The Tanglewood Symposium outlined most urgent adjustments and developments that should take place in music education. After this symposium, the participants developed the *Go Project* in 1969.[10] The goals of this project included introducing wide-ranging music programs in all schools, engaging people of all ages in learning process, and using effective resources and teaching methodologies that would help teachers be more efficient. The success of Tanglewood Symposium caused numerous changes both in research and curriculum.[11] It inspired many other symposiums, such as Ann Arbor, which also addressed the need of multiculturalism in music education.[12]

The purpose of this study is to investigate how Tanglewood Symposium influenced the nature, roles, and use of ethnic music in Silver Burdett music series textbooks. The study provides an analysis of textbooks from fourth to seventh grade published in 1962, 1968, and 1974 in three music series by Silver Burdett.

Applying a content analysis approach, the ethnic origins of 851 songs (*Music for Living*= 453, *Making Music your Own*= 249, and *Silver Burdett Music*=149) were analyzed using the SPSS statistical program. The data indicated that the music series, before and after the Tanglewood Symposium, included examples of songs from foreign cultures. Even though this symposium stressed the importance of multicultural education, it was the number of American songs[13] that increased overall while the songs of other cultures increased only for one series in 1968, falling back to the pre-Tanglewood Symposium percentage in 1974.

Despite the strong focus on American songs (46%), the *Music for Living* textbooks reflected a wide range of countries (44) in 1962. The majority of these songs were of European origin:

---

7 Stotsky, S. (1992). Academic versus ideological education in the classroom. The Education Digest Mar, 64-66.

8 Gagnon, P. (1991). Balancing multicultural and civic education will take more than social stew. The Education Digest, Dec: 7-9.

9 Mark M. L. & Gary H. L. (1992). A history of American music education. Reston, VA: Shirmer Books, a Division of Macmillan, inc.

10 Choate, R. (1968). Documentary report of the Tanglewood symposium. Washington, DC: Music Educator National Conference.

11 Mckenna, G. & Schmid, W. R. (1978). The Tanglewood syposium revisited: music in American society ten years later. Milwaukee: University of Wisconsin-Milwaukee.

12 Abeles H. F. & Hoffer C. R & Klotman, R. H. (1995). Foundations of Music Education, 2nd ed, (New York: Shirmer Books.

13 Songs that I considered in this study as an American are the ones so categorized by the authors of the textbooks.

British (8.17%), German (5.74%), French (4.42%), Czechoslovakian (2.21%), and Greek (2.21%). In addition, 4.19% of all of the songs were Indian (see Figure 1).

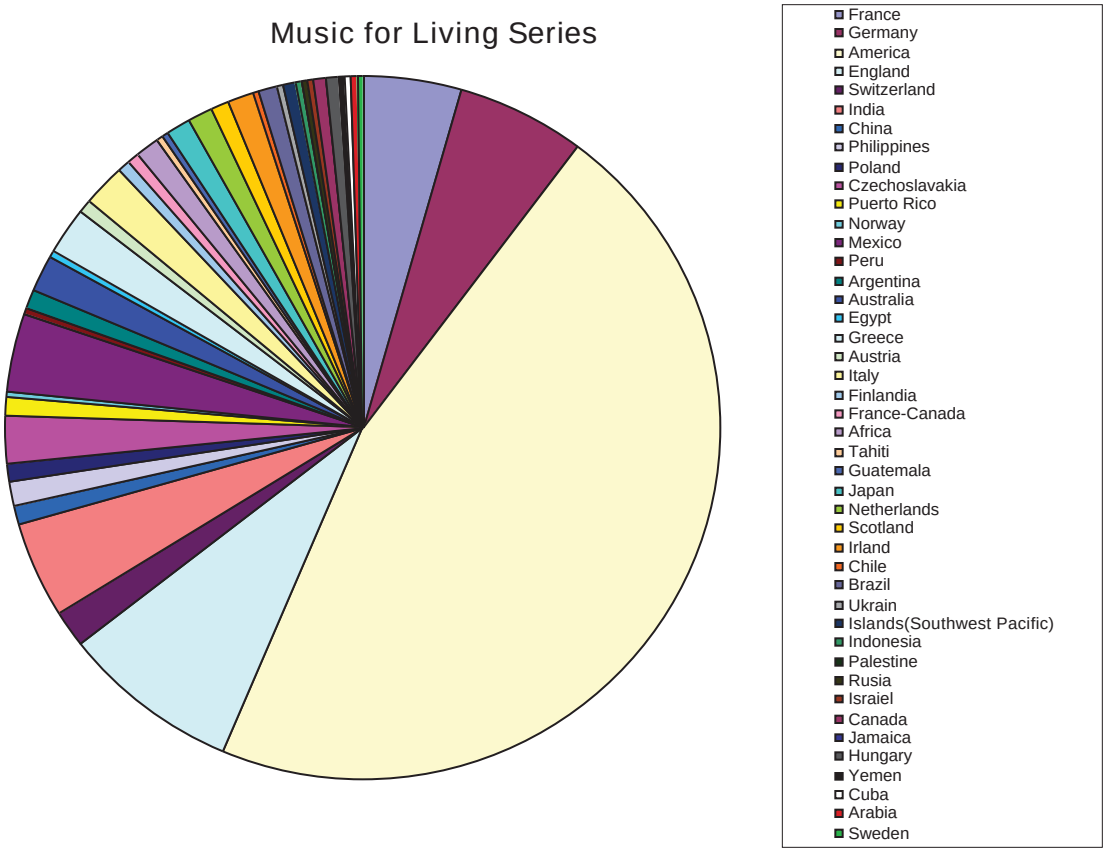


Figure 1.

In 1968, the number of American songs diminished to 30.92%. European songs, mainly German (11.24%), British (6.02%), and French (5.22%), were presented although the overall number of European songs decreased in *Making Music Your Own* series. Of the 37 countries included, eight were represented by only one song each (Poland, Guatemala, Netherlands, Brazil, Island [Southwest Pacific], Cuba, Albania, and Korea). More attention was paid to the presentation of African songs, whose number increased from 1.10% to 3.61%. The amount of Mexican songs decreased from 3.75% to 2.81% (see Figure 2).

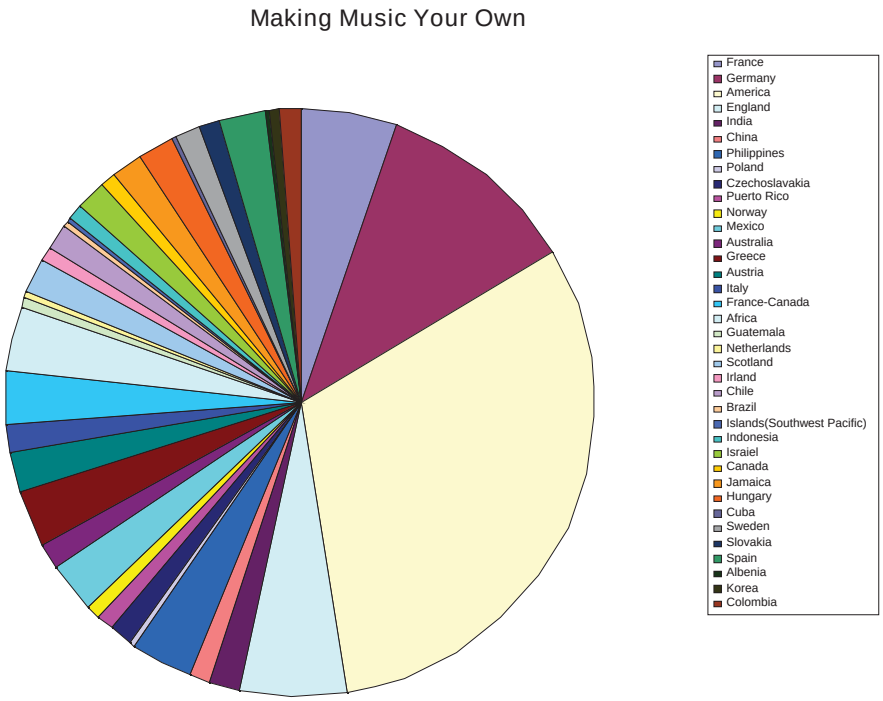


Figure 2.

According to the data, 47.97% of songs presented in the 1974 *Silver Burdett Music* series were American. Interestingly, in 1968, *Making Music Your Own* textbooks presented a broader spectrum of folk music of foreign countries than *Silver Burdett Music* in 1974 even though *Silver Burdett Music* series were published seven years after the Tanglewood Symposium. Songs of 27 countries were included in this series, and most of them were represented by only one or two songs (see Figure 3).

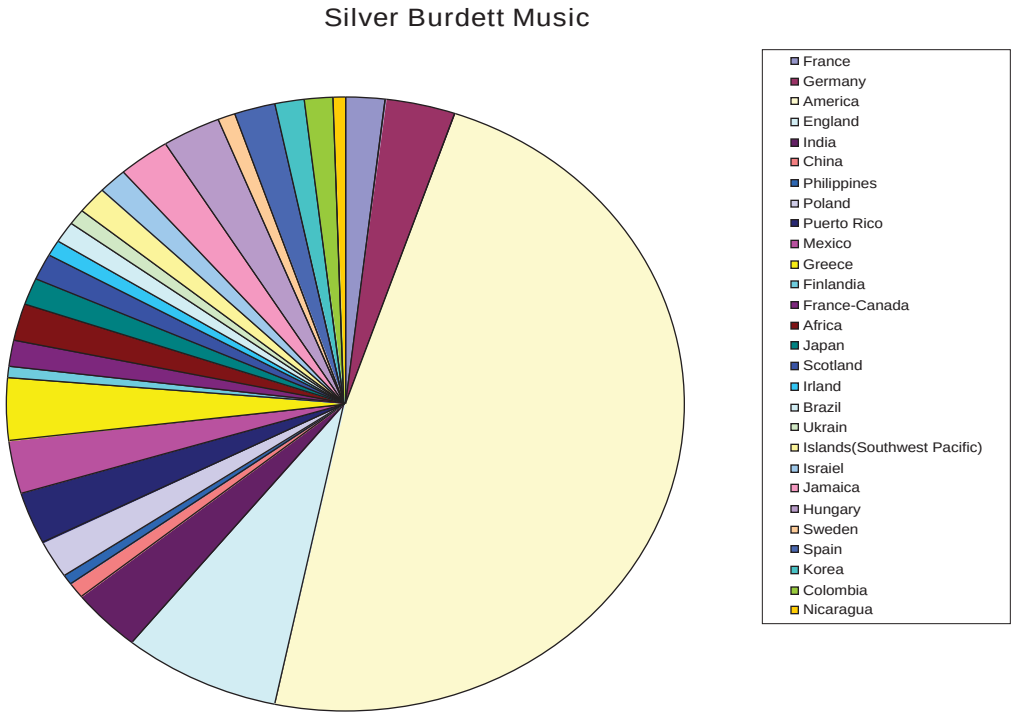


Figure 3.

Although the Tanglewood Symposium focused on the idea that multicultural education would help unite different cultures and individuals in America, it appears that the initiatives of the symposium were not accepted widely. The results imply that the notion of multiculturalism as a dividing, rather than uniting, force was prevalent at the time. The content of music education texts from that time does not reflect a conscious effort to include more music of diverse cultures; thus, the Tanglewood Symposium did not revolutionize the nature, role, or use of ethnic music in Silver Burdett music series textbooks at the time.

A multicultural society necessitates the application of an intercultural approach to education in our daily life. Today, cultural diversity and sensitivity are gaining greater attention in our profession. These issues have allowed educators to assist students in gaining culture-specific knowledge and to encourage them to be more culturally open and aware. Thus, more songs of foreign origin need to be included in today's music textbooks, and more culture-centered lesson activities should be provided. Through this cultural expansion, students may become more aware of their own cultural identity. Most importantly, while learning different folk songs, students may grow not only to appreciate their own cultural backgrounds more but also to respect and value the cultural richness of other cultures. An analysis of current music textbooks will reveal whether or not these goals are currently being realized.

References

Abeles, H. F. & Hoffer, C. R. & Klotman, R. H. (1995). *Foundations of Music Education* (2nd ed.). New York: Shirmer Books.

Banks, J. (1991). A. multicultural literacy and curriculum reform. *The Education Digest*, 10-13.

Choate, R. (1968). *Documentary report of the Tanglewood symposium*. Washington, DC: Music Educator National Conference.

Gagnon, P. (1991). Balancing multicultural and civic education will take more than social stew. *The Education Digest*, Dec: 7-9.

Gould, K. H. (1995). The misconstruing of multiculturalism: the Stanford debate and Social Work. *Social Work*, 198-204.

Mark, M. L. & Gary, H. L. (1992). *A history of American music education*. Reston, VA: Shirmer Books.

Mckenna, G. & Schmid, W. R. (1978). *The Tanglewood symposium revisited: Music in American society ten years later*. Milwaukee: University of Wisconsin-Milwaukee.

New York State Social Studies Review and Development Committee (1992). *Multicultural education benefits all students: Education in America — opposing viewpoints*. California: Greenhaven.

Ogbu, J. (1995). Understanding cultural diversity and learning. In J. A. Banks and C. A. McGee Banks (Eds.), *Handbook of research on multicultural education*.. New York: Macmillan.

Ryan, F. J. (1993). The perils of multiculturalism: Schooling for the group. *Educational Horizons*, Spring: 134-38.

Stotsky, S. (1992). Academic versus ideological education in the classroom. *The Education Digest*, March, 64-66.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Choral Directors' Self Reports of Accommodations Made for Boys' Changing Voices

Janice N. Killian
Texas Tech University

The challenges of directing choirs that include boys whose voices are in the process of changing have been well documented for decades (Swanson, 1961). Research has documented that boys' voice changes can be reliably categorized into predictable stages (Cooksey, 1977; Rutkowski, 1984), and more recent research has noted that boys' voices may change earlier than Cooksey originally determined (Moore, 1995; Killian, 1999a). The Cooksey stages and characteristics of the voice change have been disseminated widely and are generally accepted as evidenced by the contents of university choral methods books (Small & Bowers, 1997; Collins, 1999), choral textbooks for secondary schools (Killian, O'Hern & Rann, 1999) and the continuing publications addressing his ideas (Hook, 1998; White & White, 2001).

Appropriate music selection for changing voices is another challenge frequently addressed (Cooksey, 1977; Berry, 1996; Hook, 1998; White & White, 2001). While most information dealing with music selection is based on advice and testimonials, recent research has begun to address music selection from a data perspective (Rentz, 1996; Killian, 1999b; 2000; Forbes, 2001; Reames, 2001).

Although information about characteristics of the changing voice and appropriate music selection has been disseminated in both research and practitioner venues, few studies have determined whether this knowledge has affected how practitioners actually accommodate singers during the voice change. Thus the purpose of the present study is to survey choral educators regarding the techniques used to accommodate changing voice male singers with limited ranges and limited flexibility. Specifically the study was designed to explore:

1. What strategies are used to accommodate boys' voices in grades 4-9;
2. What music voicings are most frequently used in grades 4-9;
3. The degree to which teachers feel challenged when working with changing voices;
4. What needs teachers perceive regarding changing voices;
5. Teacher awareness of changing voices in their classes; and
6. Whether answers to the above questions vary by geographical location, by age of students taught, or by length of teaching experience.

Method

Preliminarily, choral educators (n= 47) attending a changing voice workshop were asked to list ways in which they dealt with changing voices in their choirs. A checklist of strategies was thus generated for the present study. Note that the checklist should not be interpreted as listing recommended methods; instead, it includes only those methods indicated by two or more pilot

participants. The checklist included the following accommodation strategies and recommended choral voicings.

Accommodation Strategies

- Select treble singers only
- Rewrite selected parts
- Instruct them to sing an octave lower
- Have separate boys' and girls' choirs
- Assign boys non-singing responsibilities
- Other (please explain)

Recommended Choral Voicings

- Sing 2-Part Treble music
- Sing TTB music
- Sing 3-Part Mixed music
- Sing SAB music
- Sing SATB music

Participants were instructed to check as many strategies on the checklist as applied. The survey included demographic information about the respondent (state of residence, years' of teaching experience, grade levels taught). Because the pilot data indicated that some directors were unaware of their singers' ranges and whether their voices were changing or not, respondents were asked to list the approximate number of boys they taught in grades 4-9 and how many of them were unchanged, changing, or changed. This question was included in an effort to determine the awareness level of the respondent regarding his/her students' changing voice. Grades 4-9 were included to encompass the most common ages of male voice change (Moore, 1995; Killian, 1999a). Additionally, the survey asked the respondents to use a Likert scale of 1-5 to rate how challenging working with changing voices was and whether they felt they had enough information about changing voices. Finally, the survey concluded with an open-ended question asking teachers what was needed to make their boys' choral experience more successful.

Subjects consisted of volunteers attending workshops on changing voices, new music, and related issues. The workshops, led by the researcher, varied in length from one hour to two days and occurred over a three-year period (1998-2000). During the workshop, participants were invited to complete the survey. No records were kept of the total number of participants in each workshop, so no data are available regarding the participant response rate.

Surveys (N = 406) were collected at music industry-sponsored events and MENC state conventions in Florida (n = 43), Minnesota (n = 82), New Mexico (n = 20), North Carolina (n = 55), North Dakota (n = 93), Tennessee (n = 22), Texas (n = 62), and Utah (n = 29).

Results

Raw data consisted of frequency data regarding accommodation strategies, Likert-scale responses to opinion questions, and content analyses of free-response opportunities. In addition, various demographic data were categorized into groups for ease of analysis. Demographic data consisted of Geographical Location (8 states), Years Taught (divided into four year segments), and Teaching Venues categorized as Children (Kindergarten—grade 5), Elementary (K—grade 8), Middle School (grades 6—8), High School (grades 9—12), Secondary (grades 6—12), and All (K—grade 12). The results of the present investigation, designed to explore how practicing music educators accommodate changing voices yielded the following:

Accommodation Strategies

The respondents indicated strategies used to accommodate boys' changing voices. Overall, 337 different respondents checked specific strategies from the provided list; 69 did not respond. Respondents could make multiple selections, resulting in 568 responses.

Octave Lower = 226
Re-Write Parts = 202
Separate Boys & Girls = 85
Select Treble Singers Only = 29
Assign Non-Singing Parts = 26

It should be noted that although Selecting Only Treble Singers was checked 26 times, only 4 people made that their only choice. Additionally, only three respondents checked Assign Non-Singing Parts as their sole choice. Thus very few directors appeared to be omitting changing voice boys completely from choral experiences.

Table 1

Frequency of Accommodation Strategies for Changing Voices by Teaching Venue

	Sing Octave Lower	Re-Write Parts	Separate Boys-Girls	Select Treble Singers Only	Assign Non-Music Duties
Children (K-5) *n*=182	102	75	13	17	11
Elementary (K-8) *n* =20	12	10	0	2	0
Middle School (6-8) *n*=80	37	48	41	5	4
Secondary (6-12) *n* =40	24	21	11	1	4
High School (9-12) *n* =6	0	3	3	0	0
All (K-12) *n*=76	51	44	16	4	7
Total Responses =566 *N*=404	226	201	84	29	26

Note. * Respondents could choose more than one accommodation strategy.

Examination of the combinations of strategies selected revealed the following:

Octave Lower = 86
Octave Lower+ Re-Write Parts = 72
Re-Write Parts = 52
Re-Write Parts + Separate Boys & Girls = 25
Re-Write Parts + Separate Boys & Girls + Octave Lower = 23
Separate Boys & Girls = 17
Octave Lower + Separate Boys & Girls = 13
All other strategy combinations were chosen less than 8 times

Chi Square analysis of the frequency data revealed that Accommodation Strategies did not vary by years taught or by teaching venue. There were, however, significant differences in the overall Accommodation x Location (X^2 (140, 337)=230, p< .0001) and Accommodation x Teaching Venue (X^2 [100,335]=143 p = .0033). Separate analyses of each accommodation strategy x location revealed a significant difference on Octave Lower (X^2 [7, 406]=23, p= .0018 and on Separate Boys & Girls (X^2 7,406] = 83, p< .0001). Similarly, separate analyses of each Accommodation Strategy x Teaching Venue revealed significant differences on Octave Lower (X^2 [5,404] = 15, p= .0107) and Separate Boys & Girls (X^2 5,404] = 75, p<.0001). Examination of the Location data indicates that only in North Carolina, Florida, and Texas did the majority of respondents *not* advise singing an octave lower. Likewise, only in Texas did the majority recommend separating boys and girls. Examination of the Teaching Venue data indicated that only Middle School and High School teachers do *not* recommend singing an octave lower. Additionally, only the Children and Elementary directors do *not* recommend separating boy and girl singers. Thus, confounding on the basis of location may be occurring because North Carolina and Florida respondents were almost all teaching elementary school, whereas Texas respondents were almost all teaching middle school. All other locations indicated a variety of teaching situations.

Music Selection Recommendations

The respondents indicated their preferred music voicings for changing voice boys. Overall, 277 different respondents checked specific voicings from the provided list; 129 did not respond. Because respondents could make multiple selections, 532 responses were indicated.

2-part Treble = 194
3-part Mixed = 154
SAB = 94
TTB = 53
SATB = 37

Examination of the combinations of music voicings selected revealed:

2-part Treble = 90
2-part treble & 3-part mixed =39
2-part & 3-part mixed & SAB = 32
3-part mixed & SAB = 28
3-part mixed = 22
TTB = 15
No other combinations were mentioned more than 8 times.

Music voicing selections did not vary by years taught. There were, however, significant differences in the Music Voicings x Location (X^2 (140, 277) = 281, $p < .0001$), and Music Voicings x Teaching Venue (X^2 (100, 275) = 287, $p<.0001$). Individual analyses revealed significant differences across each voicing x venue and across each voicing x location. It would appear, however, that there is the possibility of confounding between Location and Teaching Venue because certain venues (notably Florida and North Carolina) comprised primarily elementary teachers while other venues (notably Texas) comprised middle school teachers.

The "Other" category was a free-response section allowing respondents to explain some of their answers and to express concerns and strategies not mentioned on the checklist. Overall, 104 different respondents made comments in the "other" category; 302 did not respond. Because respondents could make multiple entries, 116 different comments were made. Individual comments were categorized into the following divisions. Evidently many respondents took the free-response section as an opportunity to explain their checklist choices in greater detail because many comments paralleled the checklist categories.

Music Selection = 20
Assign various specified parts = 16
Individual unique suggestions = 16
Rewrite Parts = 14
Non-Singing Responsibilities = 11
Octave Lower = 9
Work Individually = 5
Advice = 2
Anecdotes = 22

How Challenging? Enough Information?

Respondents were asked to use a 5-point Likert scale to answer the question: How challenging is working with changing voices (very challenging = 5 and no problem = 1)? Responses ranged from 1-5 with a mean of 3.79 and standard deviation of 1.20. Kruskal-Wallis analyses of the ordinal data were completed for How challenging x Location, Teaching Venue and Years Taught. There was a significant difference in Teaching Venue (H [7,8] = 20, p= .0012) and in Location (H [7,8] = 31, p<.0001). Examination of raw data indicated that those who might be more likely to teach changing voices (grades 7-12) believed the issue was significantly more challenging than did those who indicated they taught elementary grades. The mean response for those teaching children (K-5) was 3.48 while those teaching Middle School, Secondary or High School responses ranged from 4.03-4.20. Years Taught made no significant difference.

Likewise, respondents used a 5-point Likert scale to answer the question: Do you have enough information about changing voices (not enough information = 5 and plenty of information= 1)? Responses ranged from 1-5 with a mean of 3.81 and standard deviation of 1.23. Kruskal-Wallis analyses revealed no significant differences on Enough Information x Location, Teaching Venue or Years Taught. Responses were virtually identical across *all* categories. For example, the difference in mean responses between those teaching elementary and those teaching secondary ranged from 3.78 – 3.81.

Estimation of the Number of Changing Voice Boys

Teacher estimates of how many changing voice boys they taught and which voice change stages they were in revealed that teachers appeared able to make estimations that were in keeping with what one might expect from published data (c.f. Review of Literature). These estimations may indicate the awareness teachers have of their singers' voices or may indicate that they also were aware of what the data should show. Estimations indicated that these 406 respondents taught

27,350 boys in grades 4-9, that the number of male singers declined each grade level, and that, as expected, the percentage of changing and changed voices increased each grade.

Table 2

Estimates of Unchanged, Changing and Changed Boys in Grades 4-9

Grade	Total	Unchanged		Changing		Changed		No. Choirs	Boys/ Choir
		N	%	*N*	%	*N*	%		
4	8594	7102	82.6%	577	06.7%	74	00.9%	242	35.5
5	8407	5591	66.5%	1389	16.5%	258	03.1%	252	33.4
6	5056	2748	54.4%	1389	27.5%	351	06.9%	207	24.4
7	2361	916	38.8%	958	40.6%	241	10.2%	170	13.9
8	2051	384	18.7%	858	41.8%	588	28.7%	173	11.9
9	881	45	05.1%	245	27.8%	491	55.7%	91	09.7
Total	27,350	16,786		5416		2003		1135	24.1

Note. * Discrepancies in totals indicate differences within individual teacher estimates.

What is Needed to Make Boys' Choral Experiences More Successful?

In the free response section teachers were asked, "What do you need to make your boys' choral experiences more successful?" Of the 406 respondents, 226 answered this question. Responses were categorized into the following seven categories with a total of 240 different comments. The largest free response area indicated teachers' perceived needs for more appropriate music, more time to work with boys and more information on the changing voice.

Music/Literature Needs (*n* =86)
Teacher Request for Additional Information (*n* = 44)
Time/Scheduling Issues (*n*=33)
Teacher Anecdotes (*n* = 23)
Psychological/Sociological Considerations (*n* = 21)
Teaching Techniques/Suggestions (*n* = 17)
Recruiting Issues (*n* = 16)

Demographic Information

Demographic data regarding the volunteer participants in this study revealed information about the background of teachers attending workshops on changing voice.

Who attends workshops? Analysis of teaching experience, divided into four-year increments, revealed that teachers with 3 or fewer years experience were the largest single group (22.0%). Teachers with 25 or more years of experience continue attending workshops (13.5%), but the largest number of participants indicated less than 10 years experience (53.1%).

What teaching venues are most prevalent? The largest percentage taught K-5 Children (45.0%), followed by Middle School (19.8%), K-12 All (18.8%), 6-12 Secondary (9.9%), K-8 Elementary (5.0%) and 9-12 High School (1.5%). The rather large percentage of teachers responsible for K-12 music varies greatly by geographical location with 48.8% of North Dakota teachers and 31.2% of Minnesota teachers indicating that category.

How many teachers direct choirs? Overall 84% of the respondents indicate they direct choirs. Of those who teach K-5 children, 70% have choirs. Over 95% of all other respondents direct a choir.

Discussion

There seems to be some consensus regarding accommodation strategies recommended by the 406 respondents to this study. It should be remembered, however, that this study does not address how effective a given strategy might be. It only addresses what respondents indicate they use or recommend. Additionally, the reader is cautioned in assuming that group data indicate the recommendations of the entire group. For example, by far the most frequent accommodation strategy mentioned was asking the boys to sing an octave lower. Two groups, however, did not recommend this strategy. These groups were those teaching high school (0% recommended octave lower) and those teaching middle school (only 44% recommended octave lower). The rest of the respondents varied between 60% and 67% recommending octave lower. These data do not indicate which procedure is correct. They are simply the strategies mentioned by these subjects. Also, please note that this survey forced respondents to choose from a pre-determined checklist. Thus more variation in strategies might be present, but were not accounted for by the survey questions. The possibility was emphasized by examining the free-response section ("other"). Respondents modified their checklist answer of Octave Lower with phrases such as "octave lower if possible," "octave lower when it works," "sing octave down only on selected notes in the music," octave lower depending on the song," "use octave lower when it suits them."

Further individual strategies to specific problems were suggested. "Each boy knows his danger zone where it's too low so he can't sing in tune. When it dips down there they know to do a 'Milli Vanilli.'" "Each student has his vocal range (tessitura) taped into his folder. If the part doesn't fit, they go to another part." "I exchange parts on 2-part treble with boys choir. Bass = part 1 octave low; tenor sings where alto is written," "Have them sing baritone and alto, but change throughout the song when too high or too low." Obviously these directors are listening to their boys and making individual accommodations.

Other directors were not as specific: "Have them sing falsetto." "Give them the melody." "Allow them to sing where they can." "Allow them to sing where it's comfortable even if they may be off pitch." "Mostly ignore it."

It should also be remembered that while the information gained is interesting and demonstrates the thinking of participants across a wide geographical area, the respondents were not chosen in a random manner and so any generalizations should be made with caution. This caveat is especially important given the fact that respondents were in so many different teaching situations and that location and teaching venue appear to be confounded in that some locations had only elementary teachers and some locations had only middle school teachers.

The largest number of respondents recommended 2-part treble music, perhaps reflecting the frequency of the Octave Lower strategy. Interestingly enough, the free response section indicated a tremendous appeal for more well written, range-specific music with interesting texts that speak to young male singers. A full 36% of all comments were pleas for more music, better music with better ranges, and more appropriate texts.

The second most frequent free response was a request for more information. Teachers expressed concern about not having enough experience, material, information, and technique to work successfully with this age singer. Paralleling those concerns was the perceived need to recruit more male singers. Additionally many respondents simply felt overwhelmed with the amount of work and the lack of time. Time to work with boys was a major challenge mentioned. "I teach in seven schools; there just is no time to assess where the boys are at." "They need daily singing to be consistent, but I have my kids for two 20-minute periods a week." "I need more time. I'm at three schools a day and can't meet with them after school, during lunch or before school." "Re-writing parts is best, but with 50 boys in one class, 200 total students and no assistant, I find no time to re-write parts."

A remarkable 19% of the respondents teach K-12 at their school. The vast majority of these came from Minnesota and North Dakota. The workloads appear to be staggering! Future surveys may want to address the size of classes taught since large numbers and few contact hours seemed to be a common challenge.

Consistent with previous research (Killian, 1999b; Shields, 2001), many directors indicated a concern about the psychological aspects of the young male singer. "Boys should be encouraged that it's o.k. to have voice changes." "How do you sell them on the fact that singing is not just for girls or for gay guys?" "Need strong male singer models in high school." "Make sure personally they know they are wanted and welcome; see them as individuals." "How to build up abilities and ear without tearing down confidence." "I'm having trouble getting boys in choir because they think it is feminine." "We need to get them feeling good about singing and themselves."

Many respondents recommended "enthusiasm" as a solution to almost every category. The life and enthusiasm of these directors repeatedly appeared in their writing. "Very challenging, but I love it!!!" was typical of these responses. There were, however, some more negative responses. "Changing voice is only a problem if the boy doesn't want to sing and won't try." "I don't know if the changing voice is a bigger problem or is it just that they are such lazy singers."

The fact that these 406 teachers estimated that they worked with 27,350 boys in grades 4-9 is heartening. The magnitude of those numbers is one way to contemplate the potential influence the teaching profession has on the next generation. Even though most directors were concerned that they did not have enough boys, or that they didn't have sufficient time or information to teach them well, obviously there are thousands of young men singing in their organizations. It was most inspiring to read of their enthusiasm, their thirst for more knowledge, their willingness to share their joys and frustrations, and their hopefulness for the future.

Further study might address qualitative research regarding individuals working with changing voices, development of quantitative measures of singing achievement among changing voice boys, examination of the correlation between what directors say is the most effect method and what they actually do, exploration of techniques teachers use to determine voice ranges and stages among their choir members, and further exploration of what directors across the nation are actually doing to assist young singers to make music during their voice change. It is hoped that this study will serve as a springboard to many more specific investigations of teaching boys with changing voices and help to identify best practices from among the recommended strategies.

References

Beery, L. (March,1996). Appropriate voicings for middle school choruses. *Choral Journal.* 15-20.

Collins, D.L. (1999). *Teaching choral music*. Second Edition. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Cooksey, J.M. (1977). The development of a continuing, eclectic theory for the training and cultivation of the junior high school male changing voice. *Choral Journal, 18* (2, 3, 4, 5), 5-13, 5-16, 5-15, 5-18.

Forbes, G.W. (2001). The repertoire selection practices of high school choral directors. *Journal of Research in Music Education, 49,* 102-121.

Hook, S. (1998). Changing voice and middle school music: An interview with John Cooksey and Nancy Cox. *Choral Journal, 39*(1), 21ff.

Killian, J.N. (1997). Perceptions of the voice change process: Male adult vs. adolescent musicians and non-musicians. *Journal of Research in Music Education 45* (4), 521-535.

Killian, J.N. (1999a). A description of vocal maturation among fifth and sixth grade boys. *Journal of Research in Music Education, 47*(4). 357-369.

Killian, J.N. (1999b). Characteristics of Successful Choirs: UIL vs. Non-UIL Contests. 51-56. In R. Duke (Ed.), *Texas Music Education Research 1999.* Austin, Texas: Texas Music Educators Association. Online: http://www.tmea.org/080_College/Research/Killian.html

Killian, J.N. (2000). The relationship between music selection and contest ratings: A third year report. In R. Duke (Ed), *Texas Music Education Research 2000.* Austin, Texas: Texas Music Educators Association.

Killian, J. N. O'Hern M. & Rann, L. (1998). *Essential repertoire for the young choir. Tenor-bass teacher's edition.* (E. Crocker, Ed.) Milwaukee, Wisconsin: Hal Leonard Corporation.

Moore, M.C. (1995). The adolescent male changing voice: A study of age and attitudinal comparisons. Unpublished master's thesis. Texas Woman's University.

Reames, R.R. (2001). High school choral directors' descriptions of appropriate literature for beginning high school choirs. *Journal of Research in Music Education, 49,* 122-135.

Rentz, E. (1996). An examination of choral repertoire lists in thirteen states. *Southeastern Journal of Music Education, 8,* 65-79.

Rutkowski, J. (1984). Two year results of a longitudinal study investigating the validity of Cooksey's theory for training the adolescent male voice. In M. Runfola (Ed.). *Research symposium on the male adolescent voice*. (pp. 86-96). Buffalo State University of New York.

Shields, C. (2001). Music education and mentoring as intervention for at-risk urban adolescents: Their self-perceptions, opinions and attitudes. *Journal of Research in Music Education, 49,* 273-286.

Small, A. & Bowers, J. (1997). *Strategies for teaching elementary and middle-level chorus.* Reston, Virginia: Music Educators National Conference.

Swanson, F. (1961). The proper care and feeding of changing voices. *Music Educators Journal*, 48, 63-67.

White, C.D. & White, D.K. (2001). Commonsense training for changing male voices. *Music Educators Journal, 87*(6), 39-43,53.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Comparisons of Undergraduate Music Major's Vebal and Written Behaviors During Score Preparation Tasks in Varying Musical Contexts: A Pilot Investigation

Jeremy S. Lane
The University of Louisiana

For most novices, learning how to conduct an ensemble in musical performance is a challenging endeavor. The complex nature of conducting makes it a difficult subject to learn and equally as difficult to teach. In fact, many professional conductors doubt that it can be taught at all, subscribing to Leopold Stokowski's view that, “Conductors are born, not made. No amount of academic education can make a real conductor out of someone who is not born with the necessary qualities.” (Bamberger, 1965, p. 202).

Though it may be difficult, training novices in the art of conducting is well within the abilities of proactive, capable teachers. Educational scholar and philosopher Jerome Bruner wrote, "Any subject can be taught effectively in some intellectually honest form to any [student] at any stage of development.” (Bruner, 1960, p. 33). Findings of research in music education show that many conducting skills can be taught effectively to novices, including gesture (Johnson & Fredrickson, 1995; Orzolek, 2002), aural perception skills (Hayslett, 1996), and error detection (Decarbo, 1982; Grunow, 1980).

Although many aspects of conducting can be taught effectively, time constraints force the issue of choice with regards to which of these aspects should be taught. It would be nearly impossible for sufficient coverage to be devoted to every aspect of conducting within the allotted time frame of most conducting courses at the collegiate level. Frank Battisti, former teacher and conductor of the New England Conservatory Wind Ensemble, explains the problem, stating, “In most conducting programs… [master’s degree students] have two years total… With the starts and stops between semesters… you end up with forty-eight weeks to prepare somebody to conduct. That’s not a hell of a lot of time.” (Harris, 2001, p. 81). For this reason, instructors must choose from the multitude of related skills only those most essential to effective conducting and address these as deeply as possible.

Leading educational scholars have proposed a strategy for selection and development of curricular content based on major disciplinary ways of thinking, or expertise, as demonstrated by master practitioners (Wiggins & McTighe, 1998; Gardner, 2000; Duke, 2001). Students preparing to master particular skills should approach the subject matter in ways similar to experts in the given field. In this sense, experts provide a model that students may use as a guide for systematic instruction.

This approach seems appropriate for teaching conducting at the undergraduate level. The instructional setting of higher education implies the overall purpose of the curriculum: preparing students to perform, conduct, compose, teach, write, and think about music in the same way as professionals. Therefore, it is logical to conclude that instruction designed to cultivate expertise in conducting should 1) consist of objectives based on what expert conductors do — the “habits and concepts that reflect the best contemporary thinking of the domain” (Gardner, 2000, p. 116) and 2)

correct student misconceptions and eliminate "habits and concepts… inimical to the skilled practice of a discipline…" (Gardner, 2000, p. 116).

In an effort to identify "major disciplinary ways of thinking ," the investigator reviewed numerous published sources on expertise in conducting. These sources include interviews and informal discussions, (Bamberger, 1965; Barton, 2001, October; Casey, 1993; Chesterman, 1976, 1989; Ellis, 1994, 1997, May; Harris, 2001; Hart, 1979; Knight, 2001, March; Moss, 2002, May; Wagar, 1991; Williams, 1995;). Other sources reviewed include writings by conductors (Boult, 1924; Green & Malko, 1975; Battisti & Garafolo, 1990; Hunsberger & Ernst, 1992; Battisti, 1997, April; Corporon, 1997; Schuller, 1997) and instructional material (Prausnitz, 1983; Rudolf, 1994), and research studies in which expert conductors served as subjects (Buell, 1990; Oertel, 1998; Toney, 2000; Yarbrough, 1988, 2002, April).

Not surprisingly, experts display a wide variety of contrasting opinions and beliefs on many topics related to their craft. One issue that seems to be unanimous among conductors is the importance and necessity of consistent, thorough, systematic score study. Most experts assert that intensive score study is the first step in preparation of any piece of music and is a step that must occur well in advance of commencement of rehearsals. Knowledge gained through score study provides the basis for musical interpretation, gesture, rehearsal planning, assessment, and evaluation.

Apart from being identified as an essential element of the conductor's craft, score study skills are generally described by experts as considerably more difficult to learn than skills of conducting technique. Noted conductor Edo de Waart recalls that during his first years as a professional, trying to learn scores was a source of great frustration. "I still remember [asking conductor George] Szell… 'How do I learn it? How do you learn a score?' And I really have never gotten an answer from anybody, because nobody really ever knows." (Hart, 1979, p. 209).

Experts' descriptions of procedures reveal major disciplinary ways of thinking concerning score study that could serve as curricular foundations of beginning and intermediate conducting classes. Textbooks written for use in these classes, however, focus primarily on acquisition of physical skills necessary for beat patterns, cueing, tempo, and style; aspects of score study are given little attention (Covington, 1993; Stalter, 1996; Harris, 2001). In most cases, score study is primarily addressed in relation to aspects similar to those identified by Stokowski: how to beat time, how to read orchestral scores, and the nature of orchestral instruments. An analysis of content coverage in terms of book pages related to the topic of score study in four well known conducting texts (Greene, 1981; Labuta, 1995; Rudolf, 1994; Hunsberger, 1997) revealed small percentages of total book pages in each devoted to score study (Lane, 2002).

Few research studies have examined score study procedures and processes. Several authors have proposed theoretical models of score study that are based on analytical methods of expert musical theorists (Lentczner, 1977; Strouse, 1987; Covington, 1993; Markoch, 1995; Stalter, 1996). Most of these models reflect a "macro-micro-macro" approach to building score study knowledge. This same approach can be identified in score study methods described by expert conductors.

A common assumption implies that score study is necessary as an aid for identification and correction of performance errors in pitch and rhythm; research suggests, however, that error detection ability is independent of score study method (Grunow, 1980; Hopkins, 1991; Hochkeppel, 1993; Crowe, 1996). Other findings suggest that students should begin score study training within simple contexts of musical elements (such as single-part scores) before progressing to more complex, multi-part scores (DeCarbo, 1982; Forsythe & Woods, 1983; Sheldon, 1998; Byo & Sheldon, 2000).

The review of extant literature suggests that there is a need for research studies that examine the nature of transfer of musical skills within score study tasks. Specifically, research is needed that analyzes 1) musical skills learned and practiced by undergraduate music majors in the familiar settings of instrumental performance and applied lesson instruction and 2) the effectiveness of transfer of these skills into unfamiliar contexts related to conducting.

A qualitative research approach was determined to be appropriate for this investigation. Qualitative methodology is helpful for identifying component variables within a large scope of

behaviors and is effective in holistic analyses of processes and procedures. Merriam notes that qualitative methodology is also useful when "… a lack of theory or an existing theory fails to adequately explain a phenomenon." (2002, p. 5).

The current project is one of two pilot studies designed to examine various modes of qualitative methodology in research investigations into score study procedures. The first study (Lane, 2002) examined methods of oral interview data collection. The purposes of this investigation were to 1) examine two contrasting qualitative data collection methods and their effectiveness when used during research activities involving score study and 2) conduct an initial analysis of score study procedures used by undergraduate music majors in order to identify possible themes or trends that might warrant deeper exploration in a future research project of larger scope.

Method

Participants

Common in many qualitative research methodologies is the use of a process labeled *purposeful sampling* to select subjects (Merriam, 2002; Creswell, 1998). Purposeful sampling consists of the selection of a small number of subjects from a particular population or culture with regards to specific criteria determined by the nature of the research question.

The investigator determined, due to the exploratory nature of the study, that not more than three subjects should participate. In order to gain a better perspective on the consistency of data collection as evidenced in the various methodologies under question, it seemed reasonable to select subjects that were relatively similar. Criteria used for selecting subjects included gender, education level, primary instrument, musical achievement, and musical background.

Purposeful sampling resulted in the selection of three female subjects who were sophomore-level undergraduate music education majors attending a large southeastern university. Each played a woodwind instrument as their primary instrument and was identified as a high musical achiever by the conductor of her primary performing ensemble. All three subjects began study on their primary instrument in the sixth grade, had participated in middle and high school ensembles that maintained a high level of performance quality, and had achieved membership in the All-State honor band or orchestra of her home state.

Interview Sessions

Subjects participated in two separate individual interview sessions, each lasting approximately forty minutes. Subjects were asked to bring their primary instrument to the first interview session; instruments were not needed for the second session. All sessions were conducted within the same seven-day period and were videotaped for subsequent data collection and analysis.

In the first session, subjects were given an unmarked score of a chorale melody by J. S. Bach (Lake, 1938). Apart from pitch and rhythm notation, including fermatas, the only written markings on the score were the word *largo* at the beginning of the chorale and the name of the composer. Subjects were told to study the score and prepare as if she were going to perform the chorale in public. They were asked to write on the score the expressive markings that they intended to use, including an approximate metronome marking for tempo. They were given as much time as they needed to complete their preparation and were allowed as many opportunities as they desired to practice and refine their performance. Once they had finished preparing, the subjects performed their interpretations of the melody, which were recorded on audiotape.

Following the preparation and performance of the chorale, the investigator conducted an oral interview using open-ended questions in an effort to gain insight and perspective on subjects' decision-making processes. Six musical categories were open for exploration during the interview: tempo, dynamics, phrasing, fermata treatment, articulation, and 'other'. The investigator allowed the interview to progress according to topics described by the subjects. Therefore it is possible that not all musical categories were addressed.

Following the oral interview, the recording of the chorale performance was played back through standard stereo equipment. Subjects evaluated their performance using a written form developed

by the investigator and based on a form used in a published score-reading training method (Grunow & Froseth, 1980). Subjects rated their performance on five-point Likert-type scales in six categories similar to those discussed in the oral interview: tempo, dynamics, phrasing, fermata treatment, articulation, and overall performance. Space was provided for written comments in each category. Once the written form had been completed, the investigator engaged the subject in a brief discussion, during which she could elaborate or clarify on points addressed in the written evaluation.

For the second interview session, each subject was given an excerpt of a Grade 3 score for band (Carter, 1959) and five minutes of time during which she could study the music. Subjects were told to study the score as if they were going to teach the piece to an ensemble of 8th—9th grade-level musicians. They were allowed to sing through parts, conduct, or use any other method of score study they wished and were encouraged to notate on the score any markings that they might use to help formulate their interpretation of the music.

Following the five-minute study period, each subject heard a recording of a high school band performing the excerpt. During this listening, subjects were simply told to follow along in the score. Upon completion of the first listening, subjects were given an evaluation form similar to that used in the first interview session. The recorded performance of the excerpt was played again, and subjects were asked to evaluate the band's performance. Evaluations were done using a form similar to the one used in the first session and included 5-point Likert-type scales and space for written comments in six categories: tempo, dynamics, phrasing, pitch/intonation, articulation, and overall performance. They were allowed as many repeat hearings of the recording as they wished and were given as much time as they needed to complete their ratings.

Once the evaluation of the recording was complete, subjects were asked to identify in priority order the most salient areas of weakness that they heard in the band's performance. They were then told to develop a list of specific targets they might address if they were to rehearse the band and to describe at least one method or technique that they might use when rehearsing the targeted area. Subjects were allowed to list as many rehearsal targets as they wished and could re-listen to the recording as many times as necessary while formulating their lists.

Data Collection

The investigator used two separate data collection methods for each session. The first session utilized a mixture of written and verbal data sources and included short oral interviews between the investigator and subject at the end of each activity. In the second session, subjects' responses were completed entirely in written form. The only verbal interaction occurred when giving directions for the activity or when the investigator needed to ensure that the written statements were read and interpreted correctly.

Sessions were also structured to engage subjects in score preparation activities in varying musical contexts. Previous research (Byo & Sheldon, 2000; DeCarbo, 1983) has suggested that novice conductors are hindered by scores with multiple parts and that conducting students should be trained to interpret one single melodic line before moving on to more complex scores with multiple parts. To allow comparisons of score study procedures in these two contexts, the first interview session dealt solely with single line score study; the second session used only a multiple-part score for band. This structure also facilitated comparisons between evaluations of performances (single-line and full score excerpts) and modes of preparation (self-preparation for performance and preparation to teach others).

The transcripts of the videotaped interview sessions and the written documents provided the primary source of data for this project. Each session and written document were scripted verbatim then imported into a computer application program designed to assist in the evaluation and analysis of qualitative data (ATLAS.ti, 2002). Completed interviews for all three subjects resulted in approximately 90 minutes of aural interview data and thirteen pages of written transcripts (gathered from evaluation forms, rehearsal target lists, and marked scores).

Results

The primary mode of inquiry was that of a basic interpretive qualitative study (Merriam, 2002). This mode has been identified as the most common form of qualitative research currently found in the field of education (Merriam, 2002, p. 38) and is useful for analysis of processes and identification of variables within the context of those processes.

Data were analyzed in an attempt to identify recurring themes, which is a primary function of a basic descriptive analysis. The investigator engaged in repeated analysis with the scripted data in an effort to reach data *saturation*, or the point when "… data and emerging findings … feel saturated; that is, you begin to see and hear the same things over and over again." (Merriam, 2002, p. 26).

Other recommended strategies that were employed to ensure validity and reliability of the research design included *triangulation* and *member checks* (Merriam, 2002, p. 31). Triangulation, or examination of the data from multiple viewpoints, was achieved through analysis and comparisons of all available data sources, including transcripts, written responses, recorded performances, and observations of behaviors. Member checks, rechecking and confirming responses with the subjects themselves, was used to clarify and confirm the content of responses.

The exploratory nature of the study imposed certain limitations on the amount of data collected and the subsequent analysis. Due to the relatively low amount of raw data, generalizations of results should not be assumed. No statistical tests were employed to compare data; therefore, neither cause-and-effect nor correlations can be suggested. As the results of the analysis show, however, there are some trends worth noting that could provide interesting insight to the capabilities and weaknesses of undergraduate music students and open venues for further research.

During Gloria's oral interview the investigator asked her to describe the process she used to choose the dynamics she intended to perform in her chorale interpretation.

> "Well, just overall shape phrase, of course you would want to go up an down with every phrase but, um, and also whenever you are gonna get to a fermata, I would want to, like, you know, decrescendo into it and then hold it and then at the end there was one part where it went to, it almost tonicized, you know, C, and so the I wanted to make a contrast when it started going back to the A minor so I did a contrast."

Sophie was asked to describe her treatment of fermatas in the chorale, and responded as follows:

> "Well, usually in a Bach, like the way, umm, just the fermatas over quarter notes will just be you know, umm, for a breath in the phrase, whereas, like, over a half note generally signify more of a break in the like, it's starting over a new phrase, so that's why I circled them because these are more important, but with this one, I felt like the phrase kept going, not so much with that one, but with this one I felt the phrase kept going, so I put an arrow to make sure I didn't take too much time with that."

These two examples reflect a consistent trend that emerged when analyzing the data gathered from interview transcripts. Verbal responses tended to demonstrate an unclear, somewhat random train of thought, usually contained several incomplete sentences strung together as if they were one big sentence, and answers to the questions posed by the investigator did not emerge as a coherent, understandable unit. Also noted were consistent usage of vague statements that did not yield any specific information (i.e. 'I felt the phrase kept going').

Comparisons of verbal responses with written responses, however, provide an interesting contrast. In her evaluation of the recording of the high school band, Gloria wrote, "Pitch was OK in each section, but across the band as a whole needs work. Perhaps fixing balance issues is the best way to address this." Sophie's written response when describing possible methods she might employ to correct pitch problems reads:

> "Tune all low instrument chords slowly and then play in tempo, fixing individual problems as needed. Tune flute, oboe, cornet, and all melody instruments slowly and then in tempo, fixing individual problems as needed. Play passage from

(rehearsal letter) A to B, encouraging students to listen and to keep the intonation as it was just fixed."

Differences between the examples of verbal and written responses are immediately apparent. Analysis of the data revealed a consistent trend suggesting that written responses tended to be much shorter, more correct in terms of grammar and proper use of language, and more likely to contain specific, direct statements than were verbal responses.

When the trends emerged suggesting differences between verbal and written responses, data were reanalyzed in order to code and categorize responses. First, each response was categorized as either a *complete sentence* (reflecting correct grammar and usage of language) or an *incomplete sentence* (indicating incorrect grammar or sentence fragments.

Table 1 lists the number of complete sentences versus incomplete sentences for each subject and a grand total for the entire analysis.

Table 1
Complete/Incomplete Sentences by Subject

Participant	Complete Sentences Written	Verbal	(Total)	Incomplete Sentences Written	Verbal	(Total)
Sophie	18	0	(18)	6	5	(11)
Gloria	11	6	(17)	2	10	(12)
Maria	11	3	(14)	11	4	(15)
Total	40	9	(49)	19	19	(38)

As one can see, the discrepancy between complete and incomplete sentences in the written response modes is very pronounced. This discrepancy is not as immediately noticeable in the grand totals of the verbal responses; the differences between Gloria (2 written and 10 verbal) and Marie (11 written and 4 verbal), however, should be noted.

Data were then reanalyzed and coded according to the content of each response. Two categories of response content were used: *specific directions*, meaning that another person could read the sentence, follow the directions or actions described, and demonstrate them in observable behaviors, and *vague statements*, meaning that the content of the sentence could not be described or interpreted in specific behavioral means.

Table 2 lists the number of sentences with specific directions versus sentences with vague statements for each subject, and a grand total for the entire analysis.

Table 2
Specific Directions and Vague Statements by Subject

	Specific Directions:			Vague Statements:		
Participants	Written	Verbal	(Total)	Written	Verbal	(Total)
Sophie	16	2	(18)	10	9	19
Gloria	13	8	(21)	18	25	43
Maria	13	6	(19)	15	11	26
Total	42	16	(58)	43	45	(88)

The discrepancy between written and verbal responses containing specific directions is similar to that demonstrated earlier between complete and incomplete sentences. One should notice that differences between individual subjects in the vague statements category are not as extreme as those reported in the incomplete sentence category reported in the previous table. The large overall number of responses containing vague statements (over 60% of the total number of responses categorized) should also be noted.

Data were again reanalyzed in order to identify errors or misconceptions reflected in written and verbal responses. For example, Sophie's correlation between rhythmic values and length of fermata duration was determined by the investigator to be applied incorrectly within the context of Bach chorales. The investigator reached this decision by consulting the writings of a noted expert in music interpretation (Cone, 1968). Analysis revealed only four misconceptions that could be identified; all four occurred in verbal responses, occurred in responses categorized as incomplete sentences, and occurred when the subject was responding with vague statements rather than specific directions.

Discussion and Conclusions

Exact causes for differences between response categories reported in Tables 1 and 2 are difficult to pinpoint. It may be that writing skills are cultivated and developed during schooling more than are verbal response skills, whereas the individual differences between subjects could suggest differences in learning styles or varying strengths and weaknesses among types of intelligences, as proposed by Gardner (1983). The context of the study may also have influenced responses. The somewhat casual setting of an individual interview may encourage different categories of responses; a live teaching situation involving young students in ensemble rehearsal may encourage responses that are entirely different. The nature and scope of this study, however, do not allow for more than mere speculation.

That being stated, it would seem that further research is necessary to examine the causes for the discrepancies between verbal and written response modes with regards to clarity and preciseness of content. The high occurrence of incomplete sentences and vague statements in subjects' verbal responses is particularly noteworthy, especially when one considers that previous research has indicated that during ensemble rehearsals, high percentages of time are spent engaging in verbal behaviors (Goolsby, 1996, 1997; see also reviews by Grant & Drafall, 1991; Duke, 1999).

The low number of misconceptions revealed in the analysis is interesting. It is quite possible that the three subjects used in this study, representative of high achieving students, did not possess a great deal of misconceptions and were already adept at correctly applying their knowledge and skills in appropriate contexts. Also, the methodologies employed may not have allowed misconceptions to be uncovered during interviews or score study activities.

Apart from the salient issues concerning response modes, other aspects of subjects' score preparation deserve mention. While many of the technical aspects of each subject's chorale performance were excellent, elements of musical expression were noticeably lacking. All three subjects chose a tempo that was considerably faster than the marked tempo (*largo*), performed the chorale within a narrow dynamic range, and tended to demonstrate little variation or contrast in their treatment of fermatas, including the approach, sustain through the hold, release of the fermata and progression to the next phrase. These findings support previous research (DeCarbo, 1982; Byo & Sheldon, 2000) that suggests undergraduate music students should be trained to fully understand the musical possibilities inherent in a single melodic line before progressing to scores involving more complex textures.

The study yielded valuable feedback with regards to research methodology. The combination of verbal and written response modes within a single interview session was clearly superior to the method using only written responses. Not only were the quantity of data greater, but also higher quality data were obtained through the combination of verbal and written response modes. This

finding was substantiated in the second pilot study completed in conjunction with this project (Lane, 2002).

MENC: The National Association for Music Education (MENC) has identified several areas that should be addressed in future music education research efforts (MENC, 1998). Among these are issues related to the balance among subject matter expertise, methods classes, and practicum experience within music teacher education programs. The intention of the investigator is that the results of this pilot study and subsequent projects based on its findings be used to provide insight into this effort to achieve curricular balance in the field of conducting pedagogy.

References

ATLAS.ti: Visual Qualitative Data Analysis (Version 4.2) [Computer software]. (2002). Berlin: Scientific Software Development.

Bamberger, C. (1965). *The conductor's art.* New York: McGraw-Hill.

Barton, G. (2001, October). On the podium with Larry Rachlef. *The Instrumentalist, 56,* (3), 15-19.

Battisti, F.L. (1997, April). Conducting isn't easy. *The Instrumentalist, 51,* (9), 11-16.

Battisti, F.L. and Garafolo, R. (1990). *Guide to score study for the wind band conductor.* Ft. Lauderdale, FL: Meredith Music Publications.

Boult, A. (1924). *'The point of the stick': A handbook on the technique of conducting.* London: Paterson's Publications, Ltd.

Buell, D.S. (1990). Effective rehearsing with the instrumental music ensemble: A case study. (Doctoral dissertation, University of Wisconsin-Madison, 1990). *Dissertation Abstracts International,* 1150.

Bruner, J.S. (1960). *The process of education*. New York: Vintage Books.

Byo, J.L. & Sheldon, D.A. (2000). The effect of singing while listening on undergraduate music majors' ability to detect pitch and rhythm errors. *Journal of Band Research, 36,* (1), 26 - 41.

Carter, C. (1959). *Overture for Winds.* New York: Bourne Co.

Casey, J.L. (1993). *Teaching techniques and insights for instrumental music educators.* (Rev. ed.). Chicago: GIA Publications.

Chesterman R. (1989). *Conductors in conversation: Herbert von Karajan, Sir Georg Solti, Carlo Maria Giulini, Claudio Abbado, Eugene Ormandy, Riccardo Muti, James Levine.* London: Robson Books.

Chesterman R. (1976). *Conversations with conductors: Bruno Walter, Sir Adrian Boult, Leonard Bernstein, Ernest Ansermet, Otto Klemperer, Leopold Stokowski.* London: Robson Books.

Cone, E.T. (1968). *Musical form and musical performance.* New York: W.W. Norton & Co.

Corporon, E. (1997). The quantum conductor. In R. Miles (Ed.), *Teaching music through performance in band* (pp. 11-26). Chicago: GIA Publications.

Covington, D.L. (1993). A diagnostic and analytical model of score preparation: A clinical approach to musical morphology. (Doctoral dissertation, Florida State University, 1993). *Dissertation Abstracts International, 54-08A, 2785*.

Creswell, J.W. (1998). *Qualitative inquiry and research design.* Thousand Oaks, CA: Sage Publications.

Crowe, D.R. (1996). Effects of score-study style on beginning conductors' error-detection abilities. *Journal of Research in Music Education, 44,* 160-171.

DeCarbo, N.J. (1982). The effects of conducting experience and programmed materials on error-detection scores of college conducting students. *Journal of Research in Music Education, 30,* 187 - 200.

Duke, R.A. (2001). *Intelligent music teaching: Essays on the core principles of instruction.* Unpublished manuscript. University of Texas at Austin.

Duke, R.A. (1999). Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education, 143*, 1-43.

Ellis, B. (1997, May). Kenneth Bloomquist on the art of score study. *The Instrumentalist, 51,* (10), 12-15, 76.

Ellis, B. L. (1994). Selected band conductors' preparation to conduct selected band compositions. (Doctoral dissertation, University of Illinois at Urbana - Champaign, 1994). [On-line]. *Dissertation Abstracts International, 55-09A,* 2626. Abstract from: WEBSPIRS File: Dissertation Abstracts International, 1981-1986.

Forsythe, J.L. and Woods, J.R. (1983). The effects of conducting on the error detection ability of undergraduate and graduate instrumental conductors. *Contributions to Music Education, 10*, 27-31.

Gardner, H. (2000). *The disciplined mind.* New York: Penguin Books.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Goolsby, T.W. (1996). Time use in rehearsals: A comparison of experienced, novice and student teachers. *Journal of Research in Music Education, 44,* 286 – 303.

Goolsby, T.W. (1997). Verbal instruction in instrumental rehearsals: A comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45,* 21 – 40.

Grant, J.W. and Drafall, L.E. (1991). Teacher effectiveness research: A review and comparison. *Bulletin of the Council for Research in Music Education, 108,* 31-48.

Green, E.A.H. (1981). *The modern conductor*. (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Green, E.A.H. and Malko, N. (1975). *The conductor and his score*. Englewood Cliffs, NJ: Prentice-Hall.

Grunow, R.F. (1980). An investigation of the relative effectiveness of four modes of score preparation on visual-aural discrimination skills development. (Doctoral dissertation, University of Michigan, 1980). [On-line]. *Dissertation Abstracts International, 41-02A,* 580. Abstract from: WEBSPIRS File: Dissertation Abstracts International, 1861-1980.

Grunow, R.F. and Froseth, J.O. (1981). *MLR Instrumental Score Reading Program*. Chicago: GIA Publications.

Harris, F., Jr. (2001*). Conducting with feeling*. Galesville, MD: Meredith Music Publications.

Hart, P. (1979). *Conductors: A new generation.* New York: Charles Scribner's Sons.

Hayslett, D. (1996). The effect of movement-based training upon the aural acuity of conductors. *Contributions to Music Education, 23,* 7 – 18.

Hochkeppel, W.J. (1993). Systematic score study: effects of four methodologies of error detection achievement in instrumental conducting students. (Doctoral dissertation, Indiana University, 1993). *Dissertation Abstracts International, 54-08A,* 2937.

Hopkins, J.E. (1991). The effect of four approaches to score study on student conductors' ability to detect errors in the performance of choral music. (Doctoral dissertation, University of Illinois – Urbana/Champaign, 1991). *Dissertation Abstracts International, 52-11A,* 3854.

Hunsberger, D. and Ernst, R.E. (1992). *The art of conducting* (2nd ed.). New York: McGraw-Hill.

Johnson, C.M. and Fredrickson, W.E. (1995). The effect of aural commentary, written comments, and behavioral self-assessment on conductor intensity. *Journal of Band Research, 30,* (2), 27-38.

Knight, J. (2001, March). Freedom with discipline: An interview with Cristoph von Dohnanyi. *The Instrumentalist, 54,* (8), 12-17.

Labuta, J.A. (1995). *Basic conducting techniques* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Lane, J.S. (2002). *A descriptive analysis of undergraduate instrumental music education majors' score-study behaviors.* Unpublished manuscript, Louisiana State University, Baton Rouge.

Lake, M. (1938). *Sixteen chorales by J.S. Bach.* New York: G. Schirmer.

Lentczner, B. (1977). Guidelines and models of score preparation of atonal band literature. (Doctoral dissertation, Ball State University). *Dissertation Abstracts International, 30-01A,* 356.

Markoch, J.R., Jr. (1995). An approach to the musical analysis of wind-band literature based on analytical models used by wind-band specialists and music theorists. (Doctoral dissertation, Louisiana State University, 1995). *Dissertation Abstracts International, 56-11A,* 4312.

MENC: The National Association for Music Education, Music Education Research Task Force (1998). *A research agenda for music education.* Retrieved October 21, 2002 from http://www.menc.org.

Merriam, S.B. (2002). *Qualitative research in practice: Examples for discussion and analysis.* San Francisco: Jossey-Bass.

Moss, B. (2002, May). The secret of good conducting: An interview with James Croft. *The Instrumentalist, 56,* (10), 14-18.

Oertel, D.W. (1998). William Steinberg: His conducting and career. (Doctoral dissertation, The University of North Carolina at Greensboro, 1998). [On-line]. *Dissertation Abstracts International, 59-07A,* 2245. Abstract from: WEBSPIRS File: Dissertation Abstracts International, 1997-2000.

Orzolek, D.C. (2002). The effect of imagery and movement exercises on the ability of students to conduct expressively. *Journal of Band Research 37,* (2), 61-78.

Prausnitz, F. (1983). *Score and podium: A complete guide to conducting.* New York: W.W. Norton.

Rudolf, M. (1993). *The grammar of conducting: A practical study of modern baton technique* (3rd ed.). New York: G. Schirmer.

Schuller, G. (1997). *The complete conductor*. New York: Oxford University Press.

Sheldon, D.A. (1998). Effects of contextual sight-singing and aural skills training on error-detection abilities. *Journal of Research in Music Education, 46,* 384-395.

Stalter, T.J. (1996). The conductors' process model and its presentation in current conducting materials and methodologies. (Doctoral dissertation, University of Wisconsin-Madison, 1996). *Dissertation Abstracts International, 57-04A*, 1384.

Strouse, L.H. (1987). From analysis to gesture: A comprehensive approach to score preparation for the conductor. (Doctoral dissertation, Ball State University, 1987). *Dissertation Abstracts International, 48-03A*, 511.

Toney, H. Jr. (2000). Expressive ensemble conducting and performing: A qualitative case study of one conductor's practice. (Doctoral dissertation, University of Illinois at Urbana-Champaign, 2000). [On-line]. *Dissertation Abstracts International, 61-04A*, 3824. Abstract from: WEBSPIRS File: Dissertation Abstracts International, 1997-2000

Wagar, J. (1991). *Conductors in conversation.* Boston: G.K. Hall & Co.

Wiggins, G.P. and McTighe, J. (1998). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

Williams, J.E. (1998). *Rehearsing the band* (K.L. Neidig, Ed.). Cloudcroft, NM: Neidig Services.

Yarbrough, C. (1988). Content and pacing in music teaching. In P.J. Flowers (Ed.), *Current Issues in Music Education: Vol. 13. Student and Teacher Competencies: Interacting for Success.* Columbus, OH: Division of Music Education, School of Music, The Ohio State University.

Yarbrough, C. (2002, April). *Sequencing musical tasks: The teaching artistry of Robert Shaw.* Paper presented at the Music Educators National Conference, Nashville, TN.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

The Effects of Teaching Method and Musical Experience on Song-Learning Accuracy in a Newly-Learned Song

Charlotte P. Mizener
The University of Texas-Pan American

Singing and song learning occupy more time in the elementary music classroom than any other musical activity (Weinberg, 1988). Appropriately, researchers have produced a large amount of information regarding various aspects of children's singing and song learning. Singing ability, singing accuracy, and improvement of singing skills have all been topics of research for the better part of a century (Apfelstadt, 1986; Flowers & Sousa, 1988; Geringer, 1983; Goetze, 1985; Gould, 1969; Phillips, 1985; R. B. Smith, 1973; R. S. Smith,1961; Welch,1985). Another area of research into singing has been "song acquisition." Song acquisition may be viewed from at least two perspectives. One has to do with the way children develop song singing skills over time, as described by Davidson, McKernon, and Gardner (1981). Another perspective is the process by which singers learn new songs. Researchers have investigated various aspects of melodic learning, including learning melodic patterns (Jarjisian, 1981; Sinor, 1984) and learning whole songs (Bush, 1985; Klinger, 1996; Stadler Elmer, 1997; Weinberg, 1988).

This study was an extension of a previous investigation into song acquisition, the process by which singers learn new songs. The earlier study focused on singing accuracy in a newly-learned song according to teaching method and age or grade level. Results indicated that musical experience, not just age, might be a factor in singing accuracy and the song-learning process. The primary research questions included: (1) What are the relationships between song-learning accuracy and two teaching methods, the immersion approach and the echo approach? and (2) What are the relationships between song-learning accuracy and musical experience, including participation in elementary school music classes, participation in music ensembles, and current musical participation? Secondary research questions were: (1) How do the song-learning accuracy scores of the two songs compare? (2) What are the relationships between song-learning accuracy and gender?

Method

Subjects

Subjects (N = 73) were students, both music majors and non-majors, enrolled in music courses, either piano classes or elementary music methods classes, at a medium-sized state university. The subjects signed informed consent forms and were offered one excused absence during the semester for their participation in the study. There were five classes of subjects, two piano classes and three elementary music methods classes (two for interdisciplinary studies majors and one for music majors).

Procedure

Each class learned two folk songs similar in length and structure. "Tideo" (Song 1) and "Jubilee" (Song 2) (Hackett, 1998) were chosen because of their ranges and their similarity of form. They both had a range of an octave, and they were pitched so the range fell between C4 and C5 for female singers and C3 and C4 for the male subjects, where C4 is middle C. This range is considered suitable for the average untrained adult singer (Hackett 1998). The songs were six measures in length, and both were in quadruple meter. They both had a four measure verse-two measure refrain structure. Their forms were similar, a b a c d c for "Tideo" and a b a c d e for "Jubilee." Some classes learned Song 1 by immersion and Song 2 by the phrase-by-phrase method. The other classes learned the songs by the opposite methods. In addition, the order in which the subjects learned the songs was counterbalanced among the various classes.

The songs were taught by the researcher in a classroom setting. In both methods, the subjects heard the song a total of four times and sang it six times. The phrase-by-phrase procedure was as follows. The researcher introduced the song then asked the subjects to read the text, which was written on the chalkboard. The researcher then sang the entire song. The subjects subsequently were asked to pat the beat while listening to and echoing each phrase of the song, pairs of phrases, and finally the entire song again, for a total of three additional hearings and repeats. Subjects repeated the entire song twice more. For the songs learned by immersion, or a "sing-along" process, the researcher started as with the phrase-by-phrase approach. She introduced the song, the subjects read the words from the chalkboard, and then they listened to the entire song. After the initial listening, the researcher invited the subjects to sing along as soon as they could. Subjects and researcher sang the song together four times, then the subjects sang it without assistance twice more, for a total of six times, the same number of times as with the phrase-by-phrase method. The procedures were somewhat different from a typical song teaching approach in that no attempts were made to correct inaccuracies. Corrections and additional repeats would have made the number of hearings and repeats unequal among the classes.

The subjects learned both songs at one meeting, and they were recorded individually on audiocassette tape the second day after the learning session. The researcher gave a pitch and tempo cue, singing the first few words of the song and following them with "Ready, sing" in tempo and at the initial pitch of the song. If the singer did not match pitch immediately, the researcher briefly worked with him/her to increase the accuracy of the beginning pitch. After that, no attempts were made to correct the pitch. Subjects had access to a printed copy of the text of the two songs.

In addition to learning and performing the songs, subjects provided information about their musical experience and musical background. On the day that they learned the songs, they completed a questionnaire with questions about participation in school music classes at the elementary level, participation in music ensembles, formal coursework in music, and current musical participation.

The tapes were analyzed for melodic and rhythmic accuracy in the performance of the song. Melodic accuracy was determined by a combination of accurate pitches and intervals performed in the song. If the interval relationships in the singing performance matched those in the song within a semitone, even though they were at a different pitch level from that in the original key of the song, the pitches were considered correct. There were 46 pitches in "Tideo" and 36 in "Jubilee." Because of the limitations of the statistical program used in this study, the scores for correct pitches were collapsed into groups. For example, for "Tideo," 45 to 46 correct pitches were scored 10, 43 to 44 were 9, and so forth. Scores for "Jubilee" were treated similarly. Rhythmic accuracy was rated either accurate or not accurate. Crosstabulations were used to determine any relationships between variables in the study, including song learning accuracy, teaching method, and musical experience.

Results

Music majors comprised 30% (n= 22) of the subjects in this study, and non-music majors, 70% (n= 53). There were more females (n= 58, 79%) than males (n= 15, 21%). The level of singing accuracy was relatively high among these subjects. The majority started singing the songs at the given pitch. All but three (96%) started "Tideo" within a semitone of E4, and only seven (10%) did

not start “Jubilee” on or about G4. These results are interesting because the percentages are the same as those obtained in a similar previous study (Mizener, 1998). In the previous study, 96% of the subjects matched the starting pitch of a song having a lower beginning pitch (C4) accurately, whereas 90% started the song having a higher beginning pitch (A4) accurately. In the present study, when subjects did not start a song on the given pitch on the first trial and the researcher worked with them, they all did eventually match the pitch. Not all of those, however, used that same pitch to start singing the song. Something else that occurred as some of the subjects sang was that the first pitch of a measure or phrase was somewhat below pitch, but the following pitches of the phrase were accurate. Many of the inaccuracies seemed to be a result of not using the full range of the singing voice. Some subjects shifted to a lower key center in order to sing higher passages.

It was interesting to note that there were few significant differences among the variables investigated in this study. The first consideration was song-learning accuracy and teaching method. There were no differences between song-learning accuracy scores and the methods used to teach the songs, nor were there differences between the accuracy scores of the two songs.

Among the second group of variables, those related to musical experience, only one yielded a significant association with song-learning accuracy scores. Of participation in elementary school music classes, participation in performing ensembles, including type of ensemble or number of years of participation; participation in private lessons; or participation in formal music course work, none was associated with learning accuracy scores for either song. There was a significant relationship, however, between accuracy scores for the second song, “Jubilee,” and the question about current musical participation (X^2 (28, N = 73) = 51.76, $p < 0.034$, Cramer's $V = 0.396$). The more often subjects sang or played an instrument (daily, several times a month, a few times a year, less than once a year, or never), the higher their song-learning accuracy scores were. Gender was also associated with learning accuracy of the second song (X^2 (7, N = 73) = 15.35, $p < 0.032$, Cramer's $V = 0.462$). Females had higher scores than males.

Although this study yielded few significantly different relationships, it still supports some previous research and suggests some areas of study that merit further research. The fact that current music participation was associated with higher song-learning accuracy scores is in line with results of two related studies investigating the frequency and regularity of music instruction. These two studies support the idea that regular music instruction is necessary for improvement of singing accuracy, which is similar to the idea that frequent current music participation is associated with greater accuracy in learning a new song. Studies by Rutkowski (1996) and Moore, Fyk, Frega, and Brotons (1995/1996) indicated that students who had regular music instruction performed better on various singing tasks. The relationships between gender and learning accuracy of the second song are also similar to results of previous studies. Frequently the studies find no differences in singing accuracy between the genders (Cooper,1995; Welch, Sergeant, and White, 1995/1996). If results show gender differences in singing accuracy, girls sing more accurately than boys (Goetze and Horii, 1989; Green, 1990; Green, 1994).

The fact that this study showed only limited relationships between musical experience and singing accuracy suggests areas of future study. A more detailed investigation of the particular ensemble in which a subject performed, the number of years of participation in the ensemble, and how recently the participation took place might yield different results, and a similar investigation of private study might be undertaken. In addition, level of singing skill as determined by the performance of a familiar song might be compared to song-learning accuracy. As discussed above, a lower level of singing accuracy affected the performance of the song. It was difficult to know if the inaccuracies were a result of not remembering the melody or of not having the skills to reproduce the pitches accurately.

Singing and song learning are universal aspects of musical activity, and musical activity is a basic human endeavor. Music educators need knowledge of singing skills, singing accuracy, and the song-learning process in order to best provide opportunities for musical activity to all students. Results of this study indicate the need for further research into the area of song learning and singing accuracy.

References

Apfelstadt, H. (1986). Learning modality: A potential clue in the search for vocal accuracy. *Update, 4,* 4-6.

Bush, M. A. (1985). A comparison of two procedures for teaching a rote song: Parrot and reverse chaining. *Dissertation Abstracts International, 47,* 114A. (University Microfilms No. 86-05287).

Cooper, N. A. (1995). Children's singing accuracy as a function of grade level, gender, and individual versus unison singing. *Journal of Research in Music Education, 43,* 222-231.

Davidson, L., McKernon, P. & Gardner, H. (1981). The acquisition of song: A developmental approach. In *National Symposium on the Applications of Psychology to the Teaching and Learning of Music, Documentary Report of the Ann Arbor Symposium* (pp. 301- 315). Reston, VA: Music Educators National Conference.

Flowers, P. J. & Sousa, D. (1988, April). Pitch pattern accuracy, tonality and vocal range in the singing of preschool children. Paper presented at the meeting of the Music Educators National Conference Indianapolis, IN.

Geringer, J. M. (1983). The relationship of pitch-matching and pitch-discrimination abilities of preschool and fourth-grade students. *Journal of Research in Music Education, 31,* 93-99.

Goetze, M. (1985). Factors affecting accuracy in children's singing. *Dissertation Abstracts International, 46,* 2955A.

Goetze, M. & Horii, Y. (1989). A comparison of the pitch accuracy of group and individual singing in young children. *Bulletin of the Council for Research in Music Education, 99,* 57-73.

Gould, A. O. (1969). Developing specialized programs for singing in the elementary school. *Bulletin of the Council for Research in Music Education, 17,* 9-22.

Green, G. A. (1990). The effect of vocal modeling on pitch-matching accuracy of elementary school children. *Journal of Research in Music Education, 38,* 225-231.

Green, G. A. (1994). Unison versus individual singing and elementary students' vocal pitch accuracy. *Journal of Research in Music Education, 42,* 105-114.

Hackett, P. (1998). *The Melody Book* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Jarjisian, C. S. (1981). The effects of pentatonic and/or diatonic pitch patterns instruction on the rote-singing achievement of young children. *Dissertation Abstracts International, 42,* 2015A. (University Microfilms No. 81-24,581)

Klinger, R, Campbell, P .S., & Goolsby, T. (1998). Approaches to children's song acquisition: Immersion and phrase-by-phrase. *Journal of Research in Music Education, 33,* 24-34.

Moore, R. S., Fyk, J., Frega, A. L., & Brotons, M. (1995/1996). Influences of culture, age, gender and two-tone melodies on interval matching skills of children from Argentina, Poland, Spain, and the USA. *Bulletin of the Council for Research in Music Education, 127,* 127-135.

Phillips, K. H. (1985). The effects of group breath-control training on the singing ability of elementary students. *Journal of Research in Music Education, 33,* 179-191.

Rutkowski, J. (1996). The effectiveness of individual/small-group singing activities on kindergartners' use of singing voice and developmental music aptitude. *Journal of Research in Music Education, 44,* 353-368.

Sinor, E. (1984). The singing of selected tonal patterns by preschool children. *Dissertation Abstracts International, 45,* 3299A. (University Microfilms No. 85-01456)

Smith, R. B. (1961). A study of the effect of large-group vocal training on the singing ability of nursery school children. *Dissertation Abstracts , 21,* 3811. (University Microfilms No. 61-1667)

Smith, R. S. (1973). Factors related to children's in-tune singing abilities (Doctoral dissertation, University of West Virginia, 1973). *Dissertation Abstracts International, 34,* 7271A-7272A.

Stadler Elmer, S. (1997). Approaching the song acquisition process. *Bulletin of the Council for Research in Music Education, 133,* 129-135.

Weinberg, P. T. (1988). *A descriptive study of music teacher verbal behavior during the teaching of a new song.* Unpublished master's thesis, The University of Texas at Austin.

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, 85,* 238-247.

Welch, G. F., Sergeant, D. S., & White, P. J. (1995/1996). The singing competencies of five-year-old developing singers. *Bulletin of the Council for Research in Music Education, 127,* 155-162.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

A Survey of College Music Educators' Beliefs On How to Teach Songs to Young Children

Diane Persellin
Trinity University

When college music educators prepare preservice teachers to lead singing activities with children they can draw from a rich research background to inform their instruction. Many fine studies have been conducted to determine factors that contribute to greater vocal accuracy of children. Vocal modeling and teaching methodology are two of these factors that impact children's singing skills. Investigators have systematically studied teaching models and methods used for their effect on children's singing. These research studies include investigations of whether the teaching model is: female versus male (Hendley & Persellin, 1996; Sims, Moore, & Kuhn, 1982; Yarbrough, Green, Benson, & Bowers, 1991), a child's voice rather than an adult's voice (Green, 1987), a voice with vibrato rather than without (Yarbrough, Bowers, & Benson, 1992), and a human voice versus an instrument (Hermanson, 1972). The use of learning modalities to reinforce the teaching of songs has also been studied (Apfelstadt, 1984; Murphy & Persellin, 1993; Persellin, 1993; 1994).

Results of these research studies have guided college music methods educators when preparing pre-service teachers to lead singing activities with young children. No long-term research studies have been conducted, however, that examine whether teachers should *sing only for* children or *sing along with* children when teaching and reinforcing songs. Does one method of teaching songs produce more accurate singing?

Some music education textbooks that address methods for teaching children do not discuss specific teaching strategies or research related to the teaching of songs (McDonald & Simons, 1989; Phillips, 1992). Other textbooks discuss teaching strategies based on the teaching experiences of the authors without experimental justification. These strategies can most often be placed in one of two basic teaching sequences: (1) the whole-song method, in which a song is repeated in its entirety several times until children are able to sing it on their own, and (2) the phrase-by-phrase song method, in which the teacher divides a song into short phrases that the students echo (Anderson & Lawrence, 2001; Andress, 1998; Bayless & Ramsey, 1991; Bergethon, Boardman, & Montgomery, 1997; Campbell & Kassner, 2002; Choksy, Rozmajzl, & White, 1996; Young, W., 1990).

A third method of teaching songs to children is the "immersion method". While some texts equate "immersion" with the whole song method (Herrold, 2001; Rozmajzl & Boyer-White, 1996), "immersion" can be seen as being distinct from "whole song" in that the whole song method requires children to echo the entire song while immersion invites children to join the teacher as s/he repeatedly continues to sing the song. The immersion method can be seen frequently in children's own transmission of songs to each other (Campbell, 1998; Klinger, Campbell, & Goolsby, 1998; Harwood, 1987) as well as in other informal and formal music-making situations as in social clubs and religious services (Campbell, 1991).

Klinger, et al. (1998) conducted a study in which second graders in one class period were presented with two songs using sing-along method they referred to as "immersion" and a phrase-by-phrase method. Results indicated that children sang songs more accurately when they were

learned by singing along with the teacher through the immersion method than those learned through the phrase-by-phrase procedure. In a partial replication of that investigation, Barnes (1999) found students to perform both songs more accurately when they were taught using the immersion method. Gault (2000) conducted a partial replication of these studies using two songs. He expanded the treatment from one class period to 4 weeks of treatment on each song. He found that kindergarten and first grade students sang more accurately on one song but not the other when they were taught using the phrase-by-phrase method. He found no statistical difference between treatment groups for the second song.

Campbell and Scott-Kassner (2002) suggest a combination of these methods and encourage teachers to sing a song once or twice. Children are then asked to echo either part or all of the song before the teacher joins in the singing. The teacher's singing then provides light support for the children's singing. No comparison study was done to support these suggestions, however. Nor did Bennett (1999) investigate the validity of her recommendation, "to support their voices with yours rather than to sing with boisterous or raucous volume. Once children know the song, sing with a lighter, less projecting tone, then drop out of the singing so that their voices can take responsibility to 'keep the song going'."

Bergethon et al. (1997) recommend different teaching strategies for various situations. They admonish teachers, "**Do not sing with them**" (original emphasis) (p. 57). In more informal situations, they encourage teachers to "Sing to, and with, the children often, until singing becomes as much a part of the aural world as is the spoken word" (p. 55). No evidence in support of these recommendations, however, is provided.

Feierabend (1995; 2001) is also adamant about not singing with children. He refers to his "Golden Rule": "Sing <u>for</u> the class, not <u>with</u> the class" (p. 35, 1995) (<u>original emphasis).</u> Rutkowski and Trollinger (in press) concur, stating, "Do not sing with the children. It is important that you model for the children, but when it is their turn to sing do not sing with them." Hackett and Lindeman (2001) soften this admonition and advise music teachers, "As soon as they are able, children should sing independently and unaccompanied, without the help of the teacher's voice or recordings" (p. 38). Again, no systematic, analytical studies were performed to support these emphatic recommendations.

Thus, little agreement is evident in the literature. What is actually practiced? How do music educators from around the globe teach young children? In this paper, I present the techniques used by college music teacher educators and early children music leaders from a dozen countries. The purpose of this study was to assess methodologies that college music teacher educators and early childhood music leaders advocate to teach songs to young children.

Method

Surveys were sent to 57 early childhood college music educators – either past or current members of the Early Childhood Commission of the International Society of Music Education. In this survey, an inventory was used to obtain reactions to a series of statements on a Likert-type scale ranging from 5 (strongly agree) to 1 (strongly disagree). Educators were also encouraged to write their own comments about each statement. Thirty-five educators responded to the survey. Many provided thoughtful and insightful comments.

Results

Results indicated that 53% of the educators responding to this survey agreed or strongly agreed ("4" or "5") that "When teaching new songs to five-year-olds, teachers should *sing for* the children (asking children to echo the teacher) rather than *with* the children." Thirty-one percent disagreed or strongly disagreed ("1" or "2") (See Figure 1).

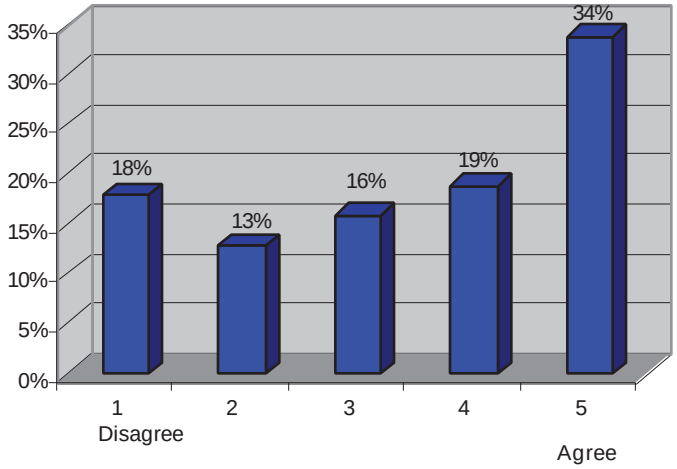


Figure 1. Responses of 35 educators to statement: "Teachers should sing for children (with children echoing) rather than with children when teaching a new song".

One teacher educator commented how she rarely sings with the children:

> "I NEVER sing with children at this age. In fact, I've actually had them stop me if I start to sing with them and say, "You're not supposed to be singing with us – it's our turn!" Well, on occasion I ask them if I can sing a song with them (once they really know the song) and they usually consent!"

Conversely, 55% agreed or strongly agreed that teachers should *sing along with* children when teaching new songs. Nineteen percent of the teachers disagreed or strongly disagreed. Most of the remaining responses (29%) were accompanied by comments stating that it would depend on the situation.

Comments about this statement included the following 3 examples:

> (1) "With five year olds, I almost never teach a song line by line with children echoing each line. I've found that, for the most part, songs for five year old should be simple enough to learn by the whole song approach. Therefore, I mostly use the "catch on method". I sing the song for them and I keep singing the whole song and encourage them to join in when they are ready."
> (2) "When teaching a new song to five-year-olds, I feel that teachers should sing with the children while they learn the song. Children should be encouraged to join the teacher's singing as soon as they are able."
> (3) "The teacher should sing with the children throughout the learning process, but listen to their performance (without singing) at the end of this process for evaluation reasons."

Seventy-two percent agreed or strongly agreed with the statement, "Teachers should sing *both with and for* children when teaching new songs". Only 12% of the respondents expressed disagreement regarding this statement. A majority of these educators indicated that it was very

important to have the flexibility to do both depending on the song, the children, and teaching context as reflected by the following statements:

> (1) "I believe that appropriate practice means that every situation is "special" and that every situation is "contextual". Applying blanket rules is not realistic and leads educators to a false sense of expected outcomes and, I believe, also limits what children can do."
> (2) "We seem to want a "one rule" to guide us in our teaching when in reality learning is not so prescriptive. Those of us who work with new teachers struggle in our responses to the begging: "Just tell me how to do it". There are times to do "echo" games, there are times to sing an entire song FOR the children, there are times to LISTEN only, and there are times to "SING TOGETHER"…. And it is through constant observations of children within the learning context that an educator can determine what is needed today with this particular group of children."

Statements that received less agreement among these educators dealt with whether teachers should or should not sing with children once a song has been learned. Forty-seven percent of the educators agreed or strongly agreed that "Teachers *should not sing with* children after a song is learned" while 35% disagreed or strongly disagreed with this statement (Figure 2).

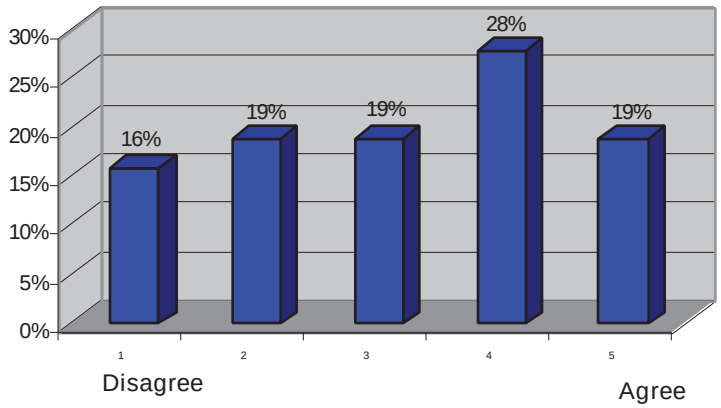


Figure 2. Responses to statement: "Teachers should not sing with children after a song is learned."

When asked to respond to the contrary statement, "Teachers *should sing with* children after a song is learned," 27% agreed or strongly agreed and 39% disagreed or strongly disagreed. Twenty-seven percent of the teachers responded with a "3" indicating that they were not comfortable agreeing or disagreeing with this question.

Many of these respondents wrote that their answers depended on many factors such as context, children, song, and singing ability of the teacher. The following comments are representative of the complexities of this issue:

> "I don't think there are general "shoulds" – it depends on the teacher's ability, the purpose of the singing in the classroom and the teacher's attitudes towards the role of music in children's lives. It also depends on the setting – day care, kindergarten, private music classes, specialist teacher, non-specialist teacher. If our first purpose is to create fine singing, and the teacher is vocally knowledgeable and capable, I think singing with the children is not as productive as singing for the children. However, if singing is regarded as an act of community, and making music as a

social, emotional, spiritual activity rather than primarily a skill building activity, then the teacher should sing with the children".

"I agree that teachers should sing with children because it is an important way to build community in the classroom. However, I also feel that as children develop confidence with a song, that opportunities for independence are important. It seems to depend on the context and the children. Also in situations in which children need more individual scaffolding according to their context and the complexity of the text/or tonal framework of the song, the teacher may need to "chime in" to enhance accuracy and success. I feel that our role is to scaffold children's singing experiences and to model active participation as well as to enhance skill development."

"I realize how strongly I've been influenced by a core belief that it is very hard for a teacher to hear how well or poorly children are singing when the teacher is singing herself. That is why I advocate not singing along with the children once a song is learned. At the same time, I realize that children learn to sing all the time by listening to tapes and CD's or singing along with an adult vocal model."

"I like to listen to the children in order to know how they are singing, but I also sing with them all the time, just to join them during the music class. I feel that is important to show the pleasure that I have to make music with them."

About half (53%) of the respondents disagreed with the statement, " I think most five-year-old children prefer to sing *with* rather than *without* a cassette or CD or recorded accompaniment." Only 8% agreed or strongly agreed with this statement.

These teachers were also asked to estimate the percentage of time spent singing with and without CD's or cassettes. The mean percentages were 20% with CD's and 80% without CD's. Teachers' comments about using CD's to accompany children's singing included:

"If children are in preschool or kindergarten classes taught by non-singing teachers, then I would far prefer that there be great vocal models on CD's for them to hear, imitate, and sing along with than off-pitch, out-of-range singing provided by some who do not know better."

"Teaching a song should not happen through a recording. The initial presentation of a song could actually be through a recording, but the actual teaching process should be done by the teacher (with no accompaniment, either recorded or live-instrumental). When children have learned the song, recorded or live-instrumental accompaniment can be used to enrich and expand their musical experiences."

"Children need to hear a good vocal model to learn to sing appropriately and that rarely occurs on recorded music. I also find these accompaniments too full and overpowering for the children."

Discussion and Conclusions

Music educators are firm in their beliefs and emphatic in their instructions to teachers on techniques best suited to teach children. But there was little agreement expressed in this survey. Only in the use of recordings was there consensus - teach without them. There was also concurrence in having the option to *sing for* as well as to *sing with* children when teaching and reinforcing songs with five-year-olds. Their comments were most insightful, indicating that one teaching model does not fit all contexts, classes or songs.

On what are their feelings based? The answers are not clear. Certainly their mentors were influential as they developed their teaching skills. Their classroom experiences were also undoubtedly important as they modified their teaching techniques according to what was working for them. The literature also likely played a role in their actions. However, there is little to no research to support their recommendations as expressed in this paper. Albeit difficult and time-consuming, studies must be done. We have initiated just such a study – one to determine whether singing for or singing with is more effective (Persellin, Klein, Smith, & Taguiam, 2002). When completed and fully analyzed, this investigation should provide some insight into this important question. Only with continued carefully controlled studies will we be able to leave the realm of intuition and hunch and be able to instruct our students using techniques based on research.

References

Anderson, W. & Lawrence, J. (2001). *Integrating music into the elementary classroom*. Belmont, CA: Wadsworth.

Andress, B. (1998). *Music for young children.* New York, NY: Harcourt Brace.

Apfelstadt, H. (1984). Effects of melodic perception instruction on pitch discrimination and vocal accuracy of kindergarten children. *Journal of Research in Music Education, 32 (1),* 15-24.

Bayless K. & Ramsey, M. (1991). *Music: A way of life for the young child* (4th ed..) New York, NY: Macmillan.

Bennett, P.D. (ed.) (1999). *Song play.* Milwaukee: WI: Hal Leonard Corporation.

Bergethon, B., Boardman, E., & Montgomery, J. (1997). *Musical growth in the elementary school* (6th ed.). Fort Worth, TX: Harcourt Brace.

Campbell, P. (1991). *Lessons from the world*. New York, NY: Schirmer Books.

Campbell, P. (1998). *Songs in their heads: Music and its meaning in children's lives.* New York, NY: Oxford.

Campbell, P.S. & Scott-Kassner, C. (1995). *Music in childhood.* New York, NY. Schirmer.

Choksy, L. (1991). *Teaching music effectively in the elementary school*. Englewood Cliffs, NJ: Prentice Hall.

Feierabend, J. (1995). *First steps in music for nursery and preschool: The curriculum*. Simsbury, CT: First Steps in Music, Inc.

Feierabend, J. (September 6-8, 2001). *First steps in music seminar*. Judson Montessori School, San Antonio, TX.

Gault, B. (2000). *The effects of pedagogical approach, presence/absence of text, and developmental music aptitude on the song performance accuracy of kindergarten and first-grade students.* Unpublished doctoral dissertation, University of Hartford.

Hackett, P. & Lindeman, C. (2001). *The musical classroom,* (5th Ed.). Upper Saddle River, NJ. Prentice Hall.

Harwood, E. (1987). *The memorized song repertoire of children in grades four and five.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Hendley, J. A., & Persellin, D.C. (1996). The comparative effects of the lower adult male voice and the male falsetto voice on children's vocal accuracy. *Update: Applications of Research in Music Education, 14 (2),* 9 – 14.

Hermanson, L.W. (1972). *An investigation of the effects of timbre on simultaneous vocal pitch acuity of young children.* (Doctoral dissertation, Teachers College, Columbia University, NY, 1971). Dissertation Abstracts International, 323, 3558A)

Herrold, R. (2001). *New approaches to elementary classroom music* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Klinger, R., Campbell, P.S. & Goolsby, T. (1998). Approaches to children's song acquisition: Immersion and phrase-by-phrase. *Journal of Research in Music Education, 46 (1),* 24-34.

McDonald, D. & Simons, G. (1989). *Musical growth and development: Birth through six*. New York, NY: Schirmer Books.

Murphy C., & Persellin, D.C. (1993). Improving vocal accuracy of first graders through learning modalities. *Texas Music Education Research,* 56-61.

Persellin, D.C. (1993). The effect of learning modalities when used in pitch matching of patterns within complete songs. *Perceptual and Motor Skills, 76 (1,)* 313-314.

Persellin, D.C. (1994). Effects of learning modalities on melodic and rhythmic retention and on vocal pitch matching by preschool children. *Perceptual and Motor Skills, 78,* 1231–1234.

Persellin, D., Klein, M., Smith, L., & Taguiam, E. (2002). The Effect of Vocal Modeling of Pitch Accuracy of Kindergarten Children. *Proceedings of the Early Childhood Commission of the International Society of Music Education*. Copenhagen, Denmark.

Phillips, K. (1992). *Teaching kids to sing.* New York, NY: Schirmer Books.

Rozmajzl, M., & Boyer-White, R. (1996). *Music fundamentals, methods, and materials for the elementary classroom teacher.* New York, NY: Longman Press.

Rutkowski, J., & Trollinger, V. (in press). Singing. In J. Flohr & R. Colwell, (Eds.), *The musical lives of young children.* Englewood Cliffs, NJ: Prentice Hall.

Sims, W., Moore, R., & Kuhn, T.L. (1982). Effects of female and male vocal stimuli, tonal pattern length and age of vocal pitch-matching abilities of young children from England and the United States. *Psychology of Music*, Special Issue: Proceedings of the IX International Seminar of Research in Music Education, 104-108.

Yarbrough, C., Bowers, J., & Benson, W. (1992). The effects of vibrato on the pitch-matching accuracy of certain and uncertain singers. *Journal of Research in Music Education, 40 (1),* 30–38.

Yarbrough, C., Green, G., Benson, W., & Bowers, J. (Winter, 1991). Inaccurate singers: An exploratory study of variables affecting pitch-matching. *Bulletin for the Council for Research in Music Education, 107*, 23-34.

Young, W. (1990). *An integrated music program for elementary school.* Englewood Cliffs, NJ: Prentice Hall.

Youngson, S.C. & Persellin, D.C. (2001). The effect of Curwen hand signs on vocal accuracy of young children. *Kodaly Envoy*, *27 (2*), 9-12.

Edited by
Mary Ellen Cavitt, The University of Texas at San Antonio

Attitudes About Performance Anxiety: Comparison of Professional and Community Orchestra Members

Beth A. Sievers and Kathy A. Thompson
University of Oklahoma

Human beings, when reacting to acute and stressful situations, trigger an in-born response that has been labeled the "flight or fight" response. Manifestations of this response can be referred to as "stage fright, performance anxiety, and musical performance anxiety" (DeForest, 1998, p. 8). Salmon (1990) defines musical performance anxiety as "the experience of persisting, distressful apprehension about and/or actual impairment of performance skills in a public context, to a degree unwarranted, given the individual's musical aptitude, training and level of preparation" (p. 2-11). Individuals with highly specialized performance skills and capabilities are affected by anxiety, and its impact has been documented in a number of settings (Salmon, 1990).

This study examines performance anxiety among both amateur and professional symphony orchestra musicians. "Stage fright, or severe anxiety about performance, is an important factor contributing to the stressfulness of a career in music, and it can seriously reduce the enjoyment experiences by amateurs and professionals playing in front of an audience" (Steptoe, 1989, p. 3-11). "Almost every public performer, regardless of the nature of his or her performance, suffers from some degree of performance anxiety" (Ely, 1991, p. 35-39).

The focus of most previous occupational research involving orchestral musicians has been related to stress. A majority of the performance anxiety research has compared professional orchestral musicians to either other professional orchestral musicians or non-music professions. Little research has been conducted comparing professional orchestral musicians' level of performance anxiety to community orchestral musicians' level of performance anxiety.

According to Ely (1991), performance anxiety has three major components: (1) physiological component – increased heart rate, sweating, shortness of breath, shaking, and other changes within the body; (2) cognitive component – our thoughts and worries about a given situation; (3) behavioral component – changing the way we think or behave to avoid anxiety-causing situations.

The objective of the present study is to help musicians and educators understand the symptoms related to performance anxiety and to learn about the effects of coping mechanisms as they relate to anxiety. Research questions that included Ely's three major components were formulated for this study. The study looked at how physiological components such as an increased heart rate, sweating, shortness of breath, and shaking affected professional and community musicians. Cognitive components of the study included how positive and/or negative comments from other musicians affected performance anxiety of a musician. Behavioral components of the study included whether musicians behaved or thought in specific ways to avoid anxiety-causing situations. Other areas of interest included comparisons of gender, positions and instrument families within the orchestra, years of performance experience beyond high school, years of study on the instrument, and primary employment.

Review of the Literature

In a study by Coons, Montello, and Perez (1995), "creative musical expression was a primary factor in helping stressed musicians transform unacceptable negative feelings into self-esteem" (p. 4-14). Steptoe (1989) compared stage fright and musicians' personal lives, and stated that one is not independent of the other. Both factors may need to be considered when developing a comprehensive approach to stress management. Steptoe recommended further research to address cognitive strategies that musicians might use when coping with stage fright.

There is evidence that performance anxiety is multifaceted and includes "fear of fear, fear of social disapproval, problems of distraction, and judgmental attitude" (Lehrer, 1987, p. 143-153). Lehrer divided the anxiety symptoms into physiological, cognitive, and behavioral components. Ely (1991) looked at the components used by Lehrer and suggested that since our experiences are shaped by our perceptions of various situations, the tendency to experience performance anxiety in public is a learned reaction. Kendrick, Cray, Lawson, and Davidson (1982) compared the results of cognitive-behavioral therapy (which involves self-instruction and focusing techniques with rehearsal) on 53 pianists who suffered from debilitating performance anxiety. The cognitive treatment program (self-instruction) and the behavior-rehearsal program (positive focus) were effective in reducing performance anxiety.

In a subjective manner, DeForest (1998) examined how a teacher's behavior and/or attitude influence(s) a student's experience of performance anxiety. Music students from a major university took the State-Trait Anxiety Inventory (Speilberger, 1968) and the Performance Anxiety Inventory (Nagel, et al. 1989). The study included qualitative aspects, such as open-ended questions, as well as quantitative instruments and analysis.

DeNelsky (1985) devised a model for the development of anxiety: (1) much early reinforcement for performance; (2) performance becomes the basis for self-esteem; (3) increasing competition among performers; (4) highly critical judges; negative feedback; (5) perfectionism becomes the goal of performance; (6) highly practiced behavior becomes less automatic; (7) the individual is more and more obsessive about a negative outcome, (unconsciously) performance anxiety mounts, and (8) a "positive feedback loop" is created, so that making a mistake is equated with failing as a human being.

Students who have not had proper guidance during critical times in their musical development might not be able to cope with anxiety and frustrations later in their careers. Hamann (1982) studied the effect of anxiety in musical performance and evaluated state and trait anxiety, curiosity, and anger under performance conditions. Results showed that years of formal training had a significant effect on the judged performance quality ($p>.05$).

Nagel, Himle, and Papsdorf (1989) developed the Performance Anxiety Inventory (PAI) to identify and compare anxiety traits. Results showed that it is possible to lower performance anxiety through a cognitive-behavioral treatment program including relaxation techniques.

A measurement of pulse rates was taken of trumpet players from a German orchestra, which revealed a mean pulse rate during performances of 91-98 beats per minute, and a maximum pulse rate of 108-115. The mean pulse rate at rest is 65-69 beats per minute (Hunsaker, 1994).

Jokinen and Kivimaki (1994) compared a group of orchestral musicians with other occupational groups and found that mental stress was high among orchestral musicians. Wind players were twice as likely as string players to suffer from mental stress. Anxiety and phobic reactions were more likely to occur in women than men; therefore, it is not surprising that more women reported performance anxiety than do men (Wesner et al., 1990). Professional musicians' level of performance anxiety lowered with age and experience (Wilson, 1997).

Method

A performance anxiety questionnaire (PAQ) was developed for this study using a 5 point Likert scale. The questionnaire consisted of 20 questions relating to performance anxiety symptoms, 16 questions relating to controlling symptoms, and demographic questions. Using the PAI (Nagel et

al., 1989) as a model, the PAQ was developed by the principal and co-principal investigators. The researchers administered the PAQ to a group of graduate students, in order to run a pilot test, and to test for face validity.

The PAQ was administered to a professional orchestra and to a community orchestra at their regular rehearsal time and place. The definition of a "professional" orchestra in this study refers to paid musicians as compared to a "community" orchestra, which does not receive payment for its services. The professional orchestra's roster is 73, and the community orchestra's roster is 63. Forty-four musicians responded in the professional group, while 59 responded in the community group. Participation was voluntary and the subjects' identities were kept anonymous.

Results

Of the 136 PAQ forms distributed, 103 forms were completed and returned (76% response rate). A performance anxiety score (PAS) was established by adding each subject's frequency responses for 17 symptoms of anxiety. The maximum score possible was 80, and the minimum score was 0. The highest actual score tabulated (72) was from a female musician. This may confirm previous research that indicated that there is a tendency for women to experience debilitating performance anxiety (Wesner et al., 1990). The lowest PAS recorded was a 0.

The most substantive frequencies included feeling confident while performing (93%), receiving positive comments from colleagues in the orchestra (86%), and an increased concentration level after making a mistake (84%). Fast heart rate was listed as a negative symptom (67%), which is comparable to Hunsaker's study (1994).

The distribution of the sample included 55% strings, 20% brass, 13% woodwinds, and 12% percussion. The mean PAS for the professional string players was 28, compared to the community string players, at 41. The string players had the greatest disparity in their PAS. The overall mean PAS was 36 for string players. It is interesting to note that the community orchestra string score is substantially higher than the community orchestra woodwind score. This is in direct contrast to the research study conducted in Finland (Jokinen & Kivimaki, 1994).

The mean PAS for the professional woodwind players was 29, compared to the community woodwind players at 33. Their overall mean score was 32. The mean PAS for the professional brass players was 46, compared to the community brass players at 44, with an overall mean score of 46. The mean PAS for the professional and community percussion players were the same at 47, with an overall mean score of 47.

The median PAS by gender were tabulated and compared between the professional and community orchestras. The women's scores for the professional orchestra (40) and the community orchestra (46) were considerably higher than the men's (31 and 36). Students' PAS were higher overall compared to music educators' and professional musicians' scores (see Table 1). Fogle (1982) found that music performance anxiety was reduced, and performance quality improved, when students were given permission to make mistakes, and when they were allowed to lower demands put on by themselves, thus reducing the "trying too hard" effect.

Table 1. *Demographics and PA median scores for sub-groups*

Demographic	Subjects *n*	PAS (all)	PAS (Professional Orch.)	PAS (Community Orch.)
Gender				
Men	46	36	31	36
Women	48	45	40	46
No response	9			
Age				
Under 25	13	46	28	47
25-35	21	35	25	43
35-44	22	42	40	45
45-54	28	37	41	37
55+	18	30	24	39
No response	1			
Employment				
Student	14	44	40	46
Educator	33	38	32	35
Performer	12	32	32	49
Other	39	31	33	39
Retired	5	32		
Section in Orchestra				
Brass	20	46	46	44
Percussion	12	47	47	47
Strings	58	36	28	41
Woodwinds	13	32	29	33
Position in Orchestra				
Principal	34	41	38	41
Section	68	36	29	41
No response	1			
Total	103	37	34	40

In general, the scores decreased as the years of study and experience increased. A chi-square was calculated comparing the professional and community orchestras to a low, medium, and high PA. Quartile deviation was used to determine the three levels, and results showed a significant difference at the ($p<.05$) level This came as a surprise since the researchers had anticipated a higher level of performance anxiety among the professional musicians.

Confidence and positive peer comments were among the important PA symptoms attitudes reported by professional and community musicians. Commonly-used coping strategies included focusing on the present rather than the past, and practicing slowly before a concert. Important coping strategies tried and reported effective by musicians included breathing techniques, visualizing a successful performance, and channeling performance anxiety to performance potential. Other coping strategies and comments reported by musicians included playing enjoyable

music before a concert, practice performing, focusing on what the music is communicating to the performer, prayer, visualizing the music, and feeling grateful for the opportunity to perform.

Conclusions and Implications

Music educators should help students realize that almost every orchestral musician experiences performance anxiety, and symptoms will probably lessen with experience and training. Although measures may be taken that alleviate particular symptoms (i.e., hand warmers), the most effective remedies appear to be those that focus attention on the music at hand and on positive images. Many subjects in this study have not tried to understand the physiological reasons for anxiety symptoms. Performers should know that trying to rid themselves of all anxious feelings might take away the potential for heightened awareness that could contribute to a better performance (Hamann, 1982).

This study cannot be generalized to all orchestral musicians because the sample size was too small. Future research might look at a larger population. Because the few positive statements in this survey have been helpful to many performers, future studies could try to determine what other coping mechanisms have been helpful to performers, thereby leading to more satisfying performance experiences.

References

Coons, E., Montello, L., & Perez, J. (1995). Confidence and denial factors affect musicians' postperformance immune response. *International Journal of Arts Medicine, 4*(1), 4-14.

Craig, K., Davidson, P., Kendrick, M. & Lawson, D. (1982). Cognitive and behavioral therapy for musical performance anxiety. *Journal of Consulting and Clinical Psychology, 50*(3), 353-362.

DeForest, G. (1998). The effects of pedagogical variables on musical performance anxiety. *Dissertation Abstracts International*, (University Microfilms No. 9901397)

DeNelsky, G. (1985). Stress of the performing arts: A special psychotherapeutic challenge. *The Third Annual Symposium on Medical Problems of Musicians*, July, 66-67.

Dews, C. & Williams, M. (1989). Student musicians' personality styles, stresses, and coping patterns. *Psychology of Music, 17*, 37-47.

Ely, M. (1991). Stop performance anxiety. *Music Educators Journal*, October, 35-39.

Fogle, D. (1982). Toward affective treatment for music performance anxiety. *Psychotherapy: Theory, Research and Practice, 19*(3), 368-375.

Hamann, D. (1982). An assessment of anxiety instrumental and vocal performances. *Journal of Research in Music Education, 50*, 77-90.

Hansaker, L. (1994). Heart rate and rhythm responses during trumpet playing. *Medical Problems of Performing Artists, 9*, 69-72.

Jokinen, M. & Kivimaki, M. (1994). Job perceptions and well-being among symphony orchestra musicians: A comparison with other occupational groups. *Medical Problems of Performing Artists, 9*(3), 73-76.

Lehrer, R. (1987). A review of the approaches to the management of tension and stage fright in music performance. *Journal of Research in Music Education, 35*(3), 143-153.

Nagel, J., Himle, D., & Papsdorf, J. (1989). Cognitive-behavioral treatment of musical performance anxiety. *Psychology of Music, 17*, 12-21.

Salmon, P. (1990). A psychological perspective on musical performance anxiety: A review of the literature. *Medical Problems of Performing Artists*, March, 2-11.

Steptoe, A. (1989). Stress, coping and stage fright in professional musicians. *Psychology of Music, 17*, 3-11.

Wesner, R., Noyes, R., & Davis, T. (1990). The occurrence of performance anxiety among musicians. *Journal of Affective Disorders, 18*, 177-185.

Wilson, G. (1997). Performance anxiety. In D.J. Hargreaves & A.C. Norton (Eds.), *The Social Psychology of Music* (pp. 229-245). New York: Oxford University Press.