

Assessing and Enhancing Feedback of Choral Conductors Through Analysis and Training

Tucker Biddlecombe
Vanderbilt University

Abstract

The purpose of this study was to determine the efficacy of a new evaluation form (Conductor Feedback Assessment Form, or CFAF) included as part of an intervention protocol designed to bring about changes in conductor feedback behaviors during choral rehearsals. Participants ($N = 32$) included professional choral directors at large high school music programs ($n = 17$) and graduate choral conducting students in their second year of study at a large Southeastern university ($n = 15$). These conductors were observed and assessed according to CFAF categories over the course of three choral rehearsals. Intermittent intervention sessions—intended to increase participants' feedback specificity—included explanation of the CFAF and videotape analyses.

Results indicated significant increases in participants' use of specific positive feedback and significant decreases in their use of nonspecific negative feedback. Teacher conductors also exhibited significant increases in their use of high magnitude positive feedback, while graduate student conductors exhibited significant increases in their use of specific negative feedback. Moreover, participant approval to disapproval ratios increased from 2:1 to 3:1 from pretest to posttest.

Most participants found the CFAF useful in increasing their feedback specificity. They reported that rehearsal pacing, choral performance, and choir members' attitudes improved when feedback became more specific. Participants also shared their personal experiences with high magnitude and anomalous feedback. Results were discussed in terms of implications for conductor training and suggestions for further research.

Keywords

choral conductors, choir rehearsing, conductor feedback assessment, specific feedback

The process of teacher feedback provides students with information about their actions in real time, encourages behaviors intended to improve their performance, and provides new information based on their resultant actions (Bandura, 1986). Every instance of feedback has a distinct characteristic. Feedback can be positive or negative, specific or general, evaluative or descriptive, verbal or nonverbal, and directed at either the individual or a group. In educational environments, the student's responsibility is to interpret and to internalize the teacher's feedback, and to make adjustments accordingly. The teacher's responsibility is to provide clear and accurate feedback, while maintaining an atmosphere that encourages and motivates students (Madsen & Madsen, 1998).

In order to measure the effectiveness of teacher feedback, researchers have most commonly measured the frequency (Kuhn, 1975; Murray, 1975; Siebenaler, 1997), specificity (Goolsby, 1997; Hendel, 1995; Speer, 1994; Yarbrough & Price, 1989), and magnitude (Whitaker, 2011; Yarbrough & Price, 1981) of feedback responses. Madsen (as cited in Grashel, 1994) suggested that positives come and go, but negatives accumulate. Some researchers have suggested that a preponderance of negative feedback may lead to discouragement, lack of achievement, and off-task behavior (Byo, 1994; Madsen & Duke, 1993), while high approval ratios improve student attentiveness (Forsythe, 1975), lessen inappropriate social behaviors (Madsen & Alley, 1979), and increase positive attitudes toward choral rehearsals (Murray, 1975). Researchers disagree, however, on the mixture of positive and negative reinforcement required for effective instruction.

In studies of effective teachers, some investigators reported more approval than disapproval feedback (Arthur, 2002; Byo, 1994; Price, 1992; Rolsten, 2011), while others found more disapprovals than approvals (Buckner, 1997; Cavitt, 2003; Derby, 2001; Duke & Simmons, 2006; Morgan, 1992; Siebenaler,

1997). Still others suggested that experienced and expert teachers often give no verbal feedback whatsoever following student responses to teacher task presentations (Brophy, 1981; Goolsby, 1997; Hendel, 1995; Yarbrough & Price, 1981).

Various authors have identified a 4:1 approval/disapproval ratio as conducive to student attentiveness and attitude ratings in classroom settings (Murray, 1975; Yarbrough & Madsen, 1998). However, many choral directors are so generous and non-selective with praise, that it carries little or no meaning to students (Bennett, 1989). Nonspecific responses of "good" or "great" following repetitions gradually lose meaning, and become habitual transition activities (Brophy, 1981). Habitual feedback may affect teacher conductors more profoundly than teachers of other subject areas, simply because of the high frequency of repetition used in ensemble rehearsals.

Yarbrough (1975) suggested the feedback history of an ensemble's interaction with a particular conductor sets up patterns of behavior that can be anticipated by students. Effective teachers modify these patterns by employing high magnitude variations in behavior, which lead to instances of high magnitude feedback that are recognizable and familiar to students. Whitaker (2011) found that students were able to recognize and identify idiosyncratic, high magnitude behaviors of their directors, including catchphrases, analogies, vocal inflections, and overt nonverbal behaviors. Behaviors identified included whispering, emphasis through voice volume, and giving strong eye contact when something went awry in rehearsal. Interestingly, directors' responses when identifying their own high magnitude feedback were nearly identical to their students' perceptions. Students rated these high magnitude instances of verbal and nonverbal behavior as very effective.

There have been many studies on conductor magnitude (Madsen, 1990, 1998; Yarbrough, 1975; Yarbrough & Price, 1981), though fewer

researchers have explored the magnitude content of approvals and disapprovals in music performance classes. Price (1983) coded a facial or verbal reinforcement as a single approval or disapproval. However, if an instance of verbal feedback was paired with an appropriate (not conflicting) facial expression, it was scored as two approvals or two disapprovals. By doubling the recorded feedback instance, the researcher attempted to reflect how intensity augmented the magnitude of feedback, thus yielding a more accurate approval/disapproval ratio for each condition.

Langer (1983) suggested that the use of positive feedback without specificity has little value when reinforcing students in academic classrooms. An overwhelming majority of research to date advocates specific feedback over nonspecific feedback in music classroom situations (Carpenter, 1988; Goolsby, 1997; Hendel, 1995; Yarbrough and Hendel, 1993; Yarbrough & Price, 1989). There is a distinct imbalance, however, in how teachers use specificity in negative and positive feedback. Some studies indicated that while specificity in negative feedback tends to be high, corresponding positive feedback tends to be less specific (Cavitt, 2003; Siebenaler, 1997; Speer, 1994). If the assertion that specific feedback is preferable to nonspecific feedback is true (Hendel, 1995; Yarbrough & Hendel, 1993; Yarbrough & Price, 1989), and that magnitude of instruction has a measurable effect on student performance in various music settings (Madsen & Duke, 1985; Yarbrough & Madsen, 1998), then perhaps the same inferences can be drawn to teacher conductor feedback in choral rehearsal settings.

Several empirical research studies were designed to improve feedback specificity through the implementation of instructional pedagogy. Price (1992) increased overall feedback specificity in pre-service teachers through intermittent training and observations. Yarbrough, Price, and Bowers (1991) achieved similar gains with experienced teachers, and

found that experienced teachers made changes in their teaching methods without the need for contingencies. Prickett (1987) significantly reduced pre-service teachers use of habitual verbal mannerisms (e.g., "OK") by using a combination of videotape observation and self-monitoring strategies. Arnold (1991) used videotaped self-analysis to increase his own use of complete teaching cycles in ensemble rehearsals. Utilizing a pretest/posttest design, he analyzed videotapes of his own teaching focusing on use of complete sequential patterns, amount of time spent on musical activities, and the specificity of feedback. Results of Arnold's self-study not only yielded increases in each of the afore-mentioned categories, but also illuminated certain realities about his own teaching, particularly the lack of specificity of his feedback to students. Arnold noted that "the lack of specificity in approvals was not only unintentional but one of which I was largely unaware. This again points out that video recording self-analysis is a valuable tool for educators. What we think we do or say is not always what we see on the tape" (p. 12).

Although considerable research exists on teacher feedback, little empirical research could be found addressing feedback specificity or feedback magnitude in choral settings. In addition, no research could be found that accounted for frequency, specificity, and magnitude in teacher feedback evaluation. Therefore, the purpose of this study was to determine the efficacy of a new evaluation form (Conductor Feedback Assessment Form, or CFAF) included as part of an intervention protocol designed to bring about changes in conductor feedback behaviors during choral rehearsals.

Specifically, the research questions for the present study were:

1. Do participants' feedback ratios of approval to disapproval increase significantly with use of the CFAF as part of an intervention protocol?
2. Does the amount of feedback specificity

by participants increase significantly with use of the CFAF as part of an intervention protocol?

3. What do participants' survey responses indicate about their perceptions of the CFAF and their experiences with high magnitude and anomalistic feedback?

Method

Development and Piloting of the Choral Feedback Assessment Form

The Choral Feedback Assessment Form (CFAF) used in this study was derived from the investigator's experiences and three published observation forms: (a) the "Choral Rehearsal

Observation Form" (Murray, 1975); (b) the "Student Teachers' Rehearsal Effectiveness Rating Scale" (Bergee, 1992); and (c) the "Teaching Unit Observation System" (Yarbrough & Price, 1981). On the CFAF, numerical values replaced common abbreviations used in previous observation forms, such as +1 replacing 'nsp' (nonspecific positive). While this change was done, in part, to clarify the scale for participants, the number scale also approximates a value hierarchy of feedback.

Figure 1 provides a sample cell of the CFAF, along with a short description of each feedback category.

Time: _____	Instruction _____					S	A	T	B	Ind.
●	-3	-2	-1	0 / R	+1	+2		+3	★	
Non-Verbal: -/ +		Cancellation: -/ +		Directive: <input type="checkbox"/>	Hustle: <input type="checkbox"/>	Verbal:				

[Star] Anomalistic positive:

Rare moment of high level shared satisfaction that most often happens in performance.

+3 = High magnitude positive

[*high affect*] "Bravo!" [*sincere*] "I'm very proud of you."

Stark change in affect, which often include idiosyncratic nonverbal behaviors.

+2 = Specific positive

"Great tone", "Right in tune that time", "Nice rhythmic energy there", "Janie – great job!"

+1 = Nonspecific positive

"Good", "Better", "Great", "That's it."

Feedback that includes no qualifiers, tend to become habitual responses functioning as a transition activity.

0 = No feedback given

"Thank you", "OK", "Again," or immediate transition to rehearsal without feedback.

-1 = Specific negative

"The f# isn't in tune altos", "tenors you're under the pitch at 16."

Feedback includes qualifiers such as specific location, problem, and/or correction strategy.

-2 = Nonspecific negative

“That’s not right altos”, “tenors you’re flat”, “you have to do better”, nonverbal negative (shaking of head, suspiration), feedback which includes no qualifiers.

-3 = High magnitude negative

[*frustrated*] “How many times to we have to go over this?” [*disappointed*] “That’s not how we do things.” Indicates high immediacy in correction, usually after multiple trial errors. Stark change in affect,, which often include idiosyncratic nonverbal behaviors.

[Black Dot] Anomalistic negative:

A moment you regret as an instructor. Walking out on a class. Singling out a student.

Non-verbal: + or – given for implied non-verbal value

Giving a ‘thumbs up’ (+1), or interpretable negative facial expression (-2), rarely specific because of the potential for multiple interpretations by students.

Directive: Disguising a negative feedback instance as an instruction

“Try it again, and this time use breath energy to keep the end of the phrase in tune,” instead of “It was flat, do it again.”

Hustle: Giving feedback during a repetition (over their singing). Cannot be verified as specific, because we can’t know everyone heard it and understood.

Figure 1. Choral Feedback Assessment Form (CFAF) sample cell with category descriptions

All feedback data were collected through videotape analysis. I conducted a pilot study to determine the initial reliability and feasibility of the CFAF as a measurement tool. I used the CFAF to score pilot participants’ videos ($N = 5$) of their choir rehearsals. Paired t-tests indicated significant differences in pilot participants’ use of categorical feedback pre and post intervention. The researcher and two trained reliability observers then independently scored 15% of the total video footage. Obtained reliability was .84 (agreements ÷ agreements + disagreements). Results of the pilot study indicated that the CFAF appeared to be a viable tool to increase participants’ use of feedback specificity, and therefore, warranted further investigation.

Participants

Participants for this study ($N = 32$) included graduate students in music education and choral conducting at a large Southeastern University ($n = 15$) and professional choral directors at high schools in New York and Florida ($n = 17$). The graduate student conductors were enrolled in a graduate-level choral conducting course, though participation was voluntary. Eight participants were males and seven were females, with teaching experience ranging from none to nine years ($M = 4.3$ years). Teacher conductor participants were recruited during a summer convention of a large professional vocal association. Eight of these professional music educators were female and nine were male, with

experience ranging from 2 years to 34 years ($M = 12$ years). Classes were videotaped with the teacher in full frontal view of the camera. Students were not videotaped at any time. Human Subjects Committee approval was granted, and school and county administrations were contacted in advance, granting permission to conduct observations at each participating school.

Procedure

This study consisted of a Time Series-Type Design, $O_1 X_1 O_2 X_2 O_3$. Each participant was observed three times teaching the same chorus class. Graduate students were observed rehearsing a lab chorus for eight-minute rehearsal periods, while the record interval for teachers included their entire rehearsal periods (minimum duration, 50 minutes), including the complete warm-up, any sight-reading practice, and all subsequent rehearsal material. When possible, teacher conductors were observed over three consecutive days, though no more than one school day elapsed between observations of teacher conductors. The time interval between graduate student observations was longer, because of a pre-existing two-week rotation schedule. Social approvals or disapprovals were not included in feedback calculation, nor were they recorded for the purposes of this study. The only requirement set forth by the researcher was that at least one piece rehearsed be at a requisite competency level to allow for a performance-oriented rehearsal to take place, rather than a note-learning rehearsal.

The first rehearsal was observed and scored as a pretest, after which the first intervention was implemented. Each intervention was conducted individually according to the schedule of the participant. The first intervention consisted of (a) thorough teaching of the CFAF categories, (b) pertinent transfers to specific events in the pretest rehearsal, and (c) setting specific feedback goals for the participating conductor.

These goals included increasing or decreasing feedback behaviors as deemed appropriate, replacing generalities with specificity, and reducing habitual feedback mannerisms. Written materials summarizing CFAF categories and corresponding goals were provided to each participant.

The second rehearsal was observed and scored, then followed by an intervention including guided videotape self-analysis, with the observer and the participant following along with the completed evaluation sheet. Participants were shown at least three examples of their classroom feedback represented on the CFAF, chosen to illustrate areas of potential improvement. A common thread for the second intervention entailed identifying habitual mannerisms (usually nonspecific reinforcement), as well as verifying instances of high magnitude positive feedback. Interventions lasted no longer than 20 minutes, and were followed by a question-and-answer period. In preparation of the third and final observation, each participant was instructed to integrate the principles of the CFAF into their teaching in the most natural manner possible.

Feedback data from third and final observation was the posttest measurement. All feedback data were collected through videotape analysis, and scored first by the investigator. Thereafter, the investigator and two trained reliability observers independently scored 15% of the total video footage. Obtained reliability was .81 (agreements \div agreements + disagreements).

Approximately one week after completing the study, participants were emailed a 10-item questionnaire intended to ascertain participant opinions regarding continued use of CFAF principles, the effectiveness of the CFAF as a teaching tool, and their experiences with high magnitude and anomalous feedback as either a teacher or a chorister. Survey return rate was 100%.

One additional factor of this study's design was the inclusion of distance-observations for

teacher conductors. Most participants were observed live. However, due to their geographical distance from the investigator, four teacher conductors were observed entirely via the Internet using video exchange. These distance-participants filmed themselves, then uploaded the videos into a Dropbox account, which were collected and scored. Interventions were conducted by FaceTime video chat. Distance participants were treated under the same conditions as live-observation participants, including durations of interventions, question-and-answer sessions, and video observation parameters.

Equipment

Rehearsal videos were recorded using a Flip Video Ultra HD video camera. Investigators and teachers viewed videos of teaching on a MacBook Pro 13.3 inch laptop computer during intervention sessions. Distance-observations were completed using two online services. Video files of rehearsals were sent through Dropbox, while interventions for distance-observations were completed using FaceTime. Folders were password-protected and were accessed only by the investigator and the participant. Questionnaires were collected using www.surveymonkey.com.

Results

Results are presented according to the research questions that guided this study. For the first and second research questions, pretest data precede the reporting of posttest results.

Research Question One: Feedback Ratios of Approval to Disapproval

During the pretest observations, mean percentages of positive to negative feedback for both graduate student conductors and teacher conductors were calculated at 67% positive to 33% negative, a ratio of approximately 2:1. This ratio was well below the 4:1 positive to negative ratio indicated as most conducive to ensemble attentiveness and attitude ratings by extant research (Kuhn, 1975; Murray, 1975; Madsen & Madsen, 1998). Calculations of positive to negative ratios excluded neutral-value feedback instances such as 'no feedback given,' and directives.

Post intervention, mean percentages of positive to negative feedback were calculated at 76% positive to 24% negative, a ratio of approximately 3:1. Tables 1 and 2 present means for teacher conductors' and graduate student conductors' use of the various CFAF categories throughout the investigation.

Table 1. *Feedback Usage Means for Teacher Conductors*

Category	<u>Observation 1</u>		<u>Observation 2</u>		<u>Observation 3</u>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High Magnitude Negative (-3)	0.06	0.24	0.06	0.24	0.06	0.24
Nonspecific Negative (-2)	4.06	2.97	2.35	2.34	1.65	1.83
Specific Negative (-1)	14.18	5.77	12.47	5.42	12.18	5.85
Directives (DIR)	11.82	5.84	10.41	5.04	9.59	4.96
Feedback Given (0)	12.94	7.58	11.82	9.57	10.53	7.62
Nonspecific Positive (+1)	25.18	15.81	19.88	7.72	17.65	5.25
Specific Positive (+2)	5.35	3.86	18.76	6.36	18.47	7.94
High Magnitude Positive (+3)	0.18	0.53	1.12	0.99	1.29	0.85

Note: Data acquired from a 50-minute rehearsal.

Table 2. *Feedback Usage Means for Graduate Student Conductors*

Category	Observation 1		Observation 2		Observation 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High Magnitude Negative (-3)	0.00	0.00	0.00	0.00	0.00	0.00
Nonspecific Negative (-2)	0.87	0.99	0.33	0.49	0.20	0.56
Specific Negative (-1)	1.07	1.16	1.67	1.59	2.87	1.92
Directives (DIR)	2.00	1.93	1.93	1.79	1.80	1.01
No Feedback Given (0)	2.53	2.07	2.40	1.68	4.07	2.09
Nonspecific Positive (+1)	6.40	4.15	5.53	4.03	6.73	5.64
Specific Positive (+2)	0.80	1.20	3.87	3.40	4.13	2.10
High Magnitude Positive (+3)	0.00	0.00	0.07	0.26	0.07	0.26

Note: Data acquired from an eight-minute rehearsal.

Research Question Two: Feedback Specificity

Pretest findings indicated that participants' negative feedback was more specific than their positive feedback. A series of paired *t*-tests revealed participant means for specific negative feedback were significantly higher than means for specific positive feedback, $t(51) = 9.57, p < .05$. When analyzing participants' overall negative feedback responses at pretest levels, findings indicated that 83% of participants' recorded negative feedback was classified as specific in nature, while 17% of participants' negative feedback was nonspecific. Conversely, only 13% of all positive feedback was classified as specific, while 87% was nonspecific. When accounting for all feedback—both positive and negative—recorded during the pretest observation, 38% of participant feedback was classified as nonspecific positive feedback, while less than 6% of positive feedback was classified as specific, making it the lowest recorded feedback CFAF category during pretest rehearsals.

Rates of specific positive feedback changed significantly post-intervention. Teacher conductors significantly increased specific positive feedback, $t(16) = 8.32, p < .05$, while significantly decreasing nonspecific positive feedback, $t(16) = 2.19, p < .05$, and nonspecific

negative feedback, $t(16) = 2.19, p < .05$, from pretest to posttest. In addition, teacher conductors exhibited a significant increase of high magnitude feedback, $t(16) = 4.15, p < .05$, in the posttest rehearsal. Ninety-three percent of all high magnitude positive responses recorded among teacher conductors occurred during the rehearsal of repertoire.

While the range of overall feedback responses varied among participants (40 to 122 instances in a 50-minute rehearsal), the average amount of feedback responses remained largely consistent from pretest ($M = 60.79$) to posttest ($M = 60.82$). Figure 2 illustrates changes in the distribution of teacher conductors' negative and positive feedback.

Graduate student conductors also showed a significant increase in specific positive feedback instances from pretest to posttest, $t(14) = 5.42, p < .05$. Graduate students exhibited significant decreases in nonspecific negative feedback, $t(14) = 2.19, p < .05$, from pretest to posttest. Unlike teacher conductors, however, graduate student conductors exhibited a significant increase in specific negative feedback, $t(14) = 4.32, p < .05$, as well as the number of overall feedback instances from pre to post, $t(14) = 2.60, p < .05$. Figure 3 shows changes in the distribution of graduate student conductors' negative and positive feedback

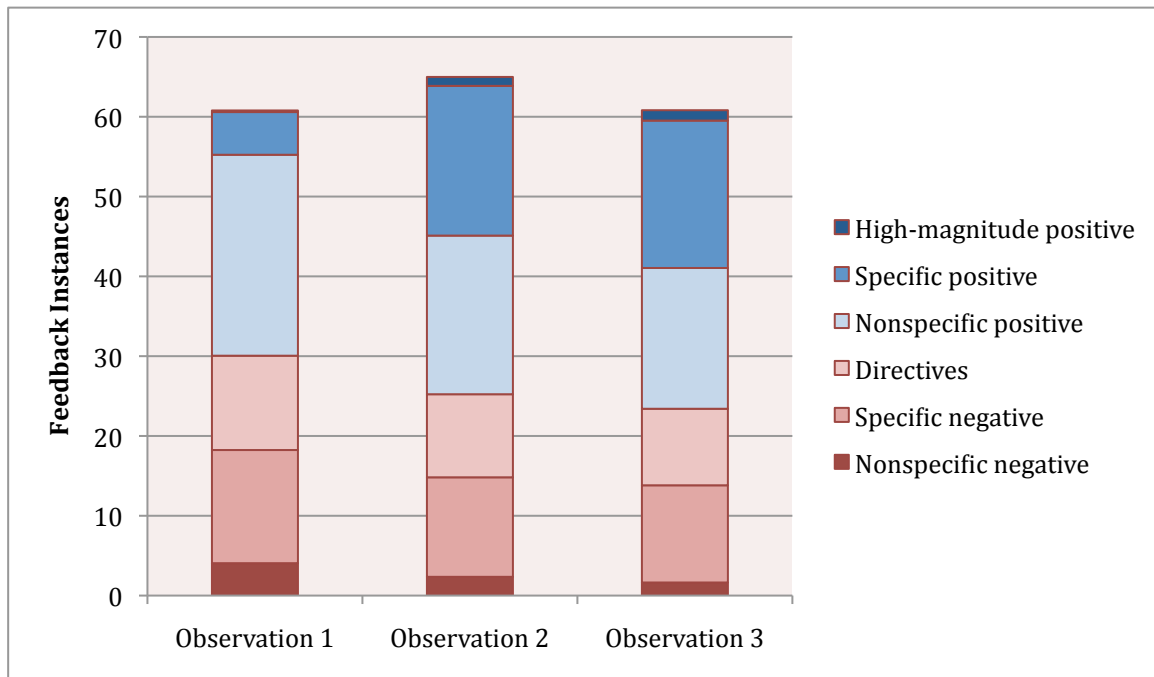


Figure 2. Changes in the distribution of feedback for teacher conductors

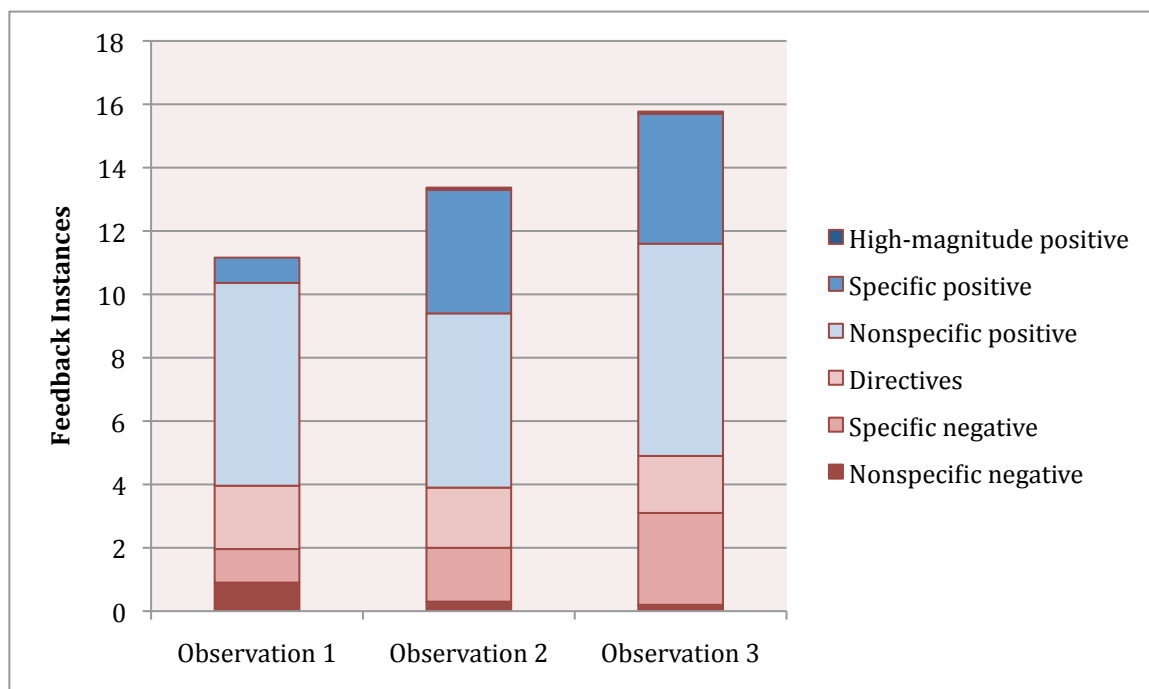


Figure 3. Changes in the distribution of feedback for graduate student conductors

Figures 4 and 5 show the significant changes in feedback usage for teacher conductors and graduate student conductors, respectively.

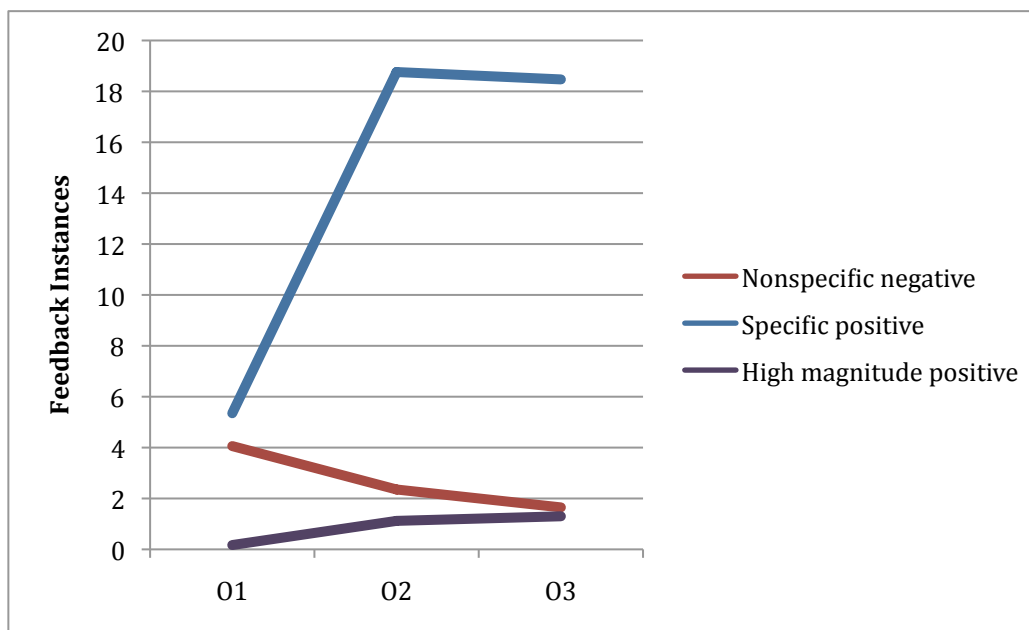


Figure 4. Significant changes in mean feedback usage for teacher conductors ($p < .05$)

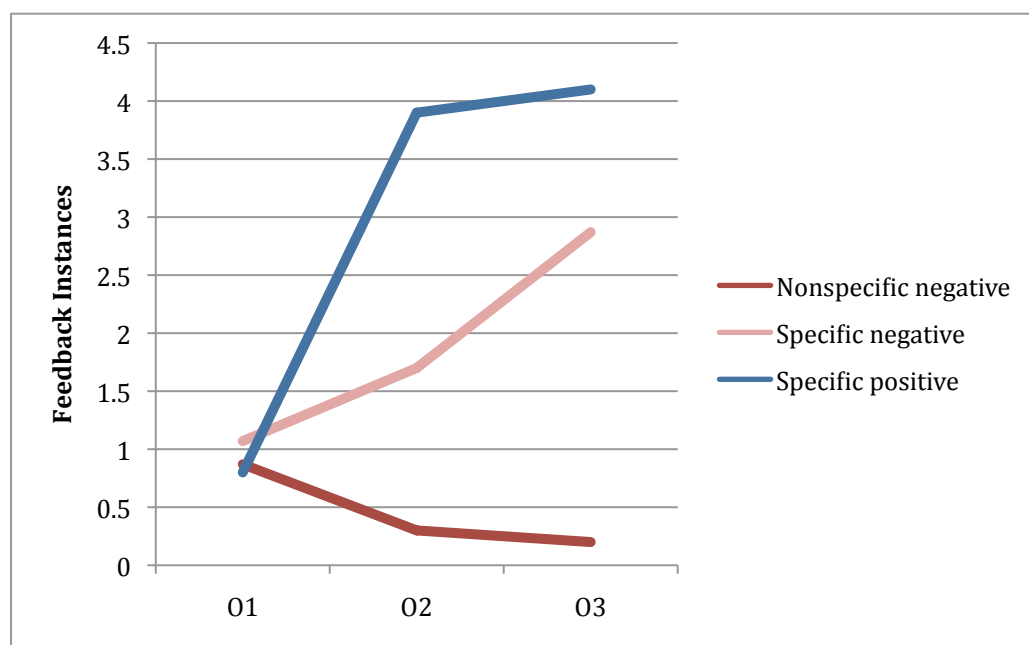


Figure 5. Significant changes in mean ratings for all graduate student conductors ($p < .05$)

Very few instances of high magnitude specific negative feedback were observed during the course of this study. Interestingly, teacher conductors participating via the Internet exhibited the only high magnitude negative feedback observed, suggesting that the presence of a live observer may have caused participants to temper their remarks. Only two instances of high magnitude specific positive feedback were observed among graduate student conductors over the course of the investigation, which may have been affected by the participation of their fellow colleagues in the lab choir, as well as the presence of their professor. There were no significant differences found in mean feedback percentages between live-observation and distance-observation participants.

Research Question Three: Participant Survey

A participant survey solicited participants' responses to seven survey statements by means of a five-point Likert-type scale, which ranged from 5 = "strongly agree" to 1 = "strongly disagree." All participants agreed that participating in the study helped to increase their feedback specificity ($M = 4.73$, $SD = 0.45$). Similarly, all participants agreed that when conductor feedback became more specific student performance improved ($M = 4.66$, $SD = 0.48$). Eighty-seven percent of respondents indicated that student attitudes improved when conductor feedback was more specific ($M = 4.36$, $SD = 0.72$). Almost three-quarters (73%) of participants thought their rehearsal pacing improved when they increased their specificity ($M = 3.83$, $SD = 0.87$). All but one respondent (96%) indicated that they maintained feedback specificity after the investigation period had ended ($M = 4.56$, $SD = 0.57$). Nearly all (93%) respondents agreed that the CFAF could be used effectively in teacher assessment and training, ($M = 4.50$, $SD = 0.73$).

The second portion of the survey consisted of three free-response prompts. First, participants

were asked to describe their use of high magnitude feedback in the classroom. These descriptions contained potential definitions of high magnitude feedback behaviors that, potentially, might assist future research. Sixty-five percent of participant responses included descriptions of high intensity, using terms such as "high energy," "proximity," "louder volume," "increased affect," "eye contact," "physical," "emphatic," and "excited." The other 35% of respondents categorized their high magnitude praise as sincere intensity, with a notable decrease in voice volume. These participants used terms such as "quiet," "goose bumps," "blissful sigh," "emotional," and "use of silence." Teacher conductors indicated idiosyncratic quotes that their students have become accustomed to because of their repeated use. A list of these quotes appears in Table 3, along with other high magnitude positive feedback examples observed and corroborated by participants during the course of this study.

The second free-response item invited participants to describe how they expressed high magnitude disapproval. One third (33%) of responses included descriptions of calmness imbued with high intensity, using terms such as "direct," "slowed speech," "low-level volume," "limited verbal," "eye contact," and "body language." Several participants indicated that they would have students put that piece of music away, accompanied by an admonition about insufficient preparation. Other responses suggested that students could identify high magnitude negative feedback through nonverbal means. Several teacher conductors referred to a common nonverbal feedback response involving strong eye contact, which they specifically referred to as "the look."

A final free-response item invited participants to describe an anomalous negative feedback moment that they had experienced as either a teacher or a student in a choral rehearsal. This moment was classified as a regrettable moment in disapproval, such as singling out a student

negatively, or leaving the classroom in frustration. A majority of respondents (72%) provided an anomalistic disapproval instance they had experienced as a student, and each description was profoundly negative. The remaining respondents either described an anomalistic moment they had experienced as a teacher (18%), or chose to leave no response

(10%). One graduate student conductor indicated that he used planned *black dot*, or anomalistic negative feedback moments, effectively in classroom settings, usually implemented before concerts. Several teacher participants wrote at length about the profoundly negative consequences of singling out a student in an ensemble class.

Table 3. *Reported and Observed Instances of High Magnitude Positive Feedback*

<i>Category</i>	<i>Quotation</i>	<i>Accompanying Nonverbal</i>
Reported Energetic	"Thank You, JESUS!"	Hands in the air
	"YES!"	Stand up from piano bench
	"Click SAVE on your mental computer,"	Point to temple
	"That's what I'm talking about!"	
	"Yes, that was AMAZING!"	Throw arms out
	"That's it!"	Fist pump
	"I'd pay money to hear that!"	Slam hand on piano
	"I'll take that every time."	
	"Yes! That's exactly what I want"	Point to a section
	"Wooooo!"	Throw arms up in the air
Reported Sincere	"You got me."	Running hands down arm (indicating Goosebumps)
	"You made beautiful music happen"	Smile of satisfaction
Observed Energetic	"Wow, that was beautiful"	Silence, strong eye contact
	"Do it twice in a row at (school) and it's a tradition"	Smile of joy
	"Do it again so I know it wasn't a fluke!"	Clap hands
	"Yeah!"	Whispered, stamp foot
Observed Sincere	"You knocked my socks off!"	Clap hands
	"I'm very proud of you"	Intense eye contact
	"I'm going to stop because that was so beautiful"	Hand on heart
	"You sound like college students"	Low voice, eye contact
	"Live in that moment for a second"	Self-embrace
	"You gave me shivers"	Self-embrace

Discussion

The results of this study indicate that conductors are significantly more specific in their negative feedback than in their positive

feedback. Negative feedback instances observed in this study often included multiple qualifiers intended to identify the characteristics of the error, such as: precise location, vocal modeling, verbal clarification, nonverbal gestures (conducting) and a suggested course of action

for students in a subsequent trial. It is presumed that conductors use these qualifiers to increase the efficiency of the correction process by clearly identifying problem areas, and fixing them with as minimal amount of repetition as possible. In addition, research has suggested that the infusion of specificity into disapproval lessens the negative magnitude of that disapproval (Duke & Henninger, 1998, 2002). The current study reflects the idea that specificity in disapproval is common practice, as participants' levels of nonspecific negative feedback were extremely low pre and post intervention.

The positive feedback that followed specific negative feedback was most often general in nature, and lacked identifying qualifiers about what was correct. Therefore, it can be assumed that conductors believe that students are carrying over the information included in the negative feedback to the ensuing positive feedback. For example, if students are instructed to "fix the F-sharp in measure 13," the teacher assumes that an ensuing response of "good" verified to students that they accomplished that goal. It would seem, however, that the carryover of qualifiers from specific negative feedback would only remain intact if every repetition were carried out with only one goal or instruction in mind. For instance, a conductor might say, "try it again, but this time add more space in the vowel on the word *fought*, and tenors get that F-sharp a little higher." Assuming the conductor affirms the ensuing repetition with nonspecific positive approval (e.g., "good!"), which aspect of the instruction was correct—the F-sharp, or the vowel quality? In addition, nonspecific praise following repetitions that had no specific instruction provides little to no information to students about what that trial accomplished, and is characterized by Yarbrough and Price (1989) as a feedback mistake on the part of the teacher.

The superficial merit of nonspecific praise is not a new notion. In fact, it is difficult to find a

pedagogical or philosophical approach that advocates the practice. Nevertheless, the current study suggests that nonspecific positive feedback is the default affirmation used by choral conductors in rehearsals. This study also reveals that conductor approval is not only overwhelmingly general, but it is also largely habitual in nature. Every participant used habitual nonspecific positive responses like "good" to varying degrees. These habitual responses were so prevalent, in fact, that reducing them was a primary intervention goal for 80% of participants in this study. Many participants noted that the most difficult aspect of increasing specificity in positive feedback was altering their habitual patterns of reinforcement. For others, simply being made aware of those habits promoted change. One participant noted, "I have heard this information before, and have aimed to provide specific feedback in my rehearsals, but this study helped to identify habits of mine, as well as give real evidence of my implementation of this technique."

During interventions, a common topic of conversation centered on conductors' prevailing dissatisfaction with students' retention of previously learned material. Both teacher conductors and graduate student conductors lamented wasted rehearsal time spent on re-teaching previously taught material. Several participants noted increased levels of student retention as their positive feedback became more specific; suggesting that lack of retention might be related to a lack of specificity in positive feedback. Nonspecific positive feedback informs the student that they did well, but specific positive feedback informs the student *what* they did well. As one participant noted: "I usually find myself re-teaching certain passages over multiple rehearsals. I find myself asking them, 'Don't you remember doing this yesterday?' and I get blank looks. As my feedback got more specific, their performance didn't necessarily improve (I still had to revisit

things), but they remembered what we had previously done, and attained high levels quicker than before.” It is possible that specific positive feedback encourages better student recall of rehearsal concepts, leading to better overall student retention.

The results of the current study also suggest that the ideal 4:1 ratio of approval to disapproval may need to be reexamined to account for feedback characteristics, such as specificity and magnitude. The ratios of positive to negative feedback observed during the current study were recorded at approximately 2:1 for pre-intervention levels, and increased to 3:1 during post-intervention observations. Only two participants were recorded at the 4:1 ratio for multiple observations, and only one participant, a graduate student, was recorded at the 4:1 ratio for all three observations. Each of these participants exhibited inordinate amounts of nonspecific praise, suggesting that the easiest way to attain the 4:1 ratio is to give increased general approval. While the 4:1 ratio was created in an effort to decrease disapprovals, that ratio would only seem to be effective when the magnitude of feedback instances is relatively equal. For instance, four nonspecific responses of “good” do not necessarily balance a high magnitude negative response of, “that was an incredibly poor effort today.”

Participants were often observed giving more verbalization than was necessary to illustrate instructional or feedback goals. Excess verbalizations exhibited most often by participants included multiple instructions before repetitions, excessive descriptions of sound ideals, repeated instructions, off-task comments, and habitual mannerisms. These superfluous verbalizations made the task of giving specific feedback more difficult, and were addressed by the investigator during interventions. In addition, participants tended to increase verbalization during feedback, often delaying the immediate delivery of feedback to students. Considerable attention was paid during intervention sessions to reducing instructions to

one concept, as well as minimizing the verbal content of feedback. These pedagogical suggestions were based on research that identifies low instructor verbalization, immediacy of feedback, and single-concept repetitions as best practice in choral rehearsals (Arthur, 2002; Napoles, 2006; Yarbrough & Price, 1989).

An effective pedagogical method for defining feedback magnitude can be gleaned from examining the numerical values represented on the CFAF. During intervention treatments, the investigator commonly referred to the numerical values when discussing feedback chaining, or multiple feedback instances given in pursuit of a musical goal. If a specific negative [-1] is given for a repetition, a specific positive [+2] should be given in response if the students have accomplished the musical goal. When added together, the remainder is ‘+1,’ suggesting that the feedback chain concludes with a positive value. If a nonspecific positive [+1] is given in response to the same repetition, the equation result is 0, and the feedback chain concludes with a neutral value. While these equations are not intended to determine the feedback balance of an entire rehearsal, they were often used during interventions to articulate feedback goals.

Another important concept of the CFAF gleaned from intervention sessions is that any category in the feedback continuum represents a potential response for a conductor. Most participants immediately increased amounts of positive feedback during the second observation, probably under the assumption that the goal of the study was to increase positivity. During interventions, participants tended to view the positive side of the scale as successful teaching, and the negative side as the area to be avoided. The investigator was careful to point out that every numerical category in the CFAF can be effectively used if contingently applied during instruction. The analogy used with every participant referred to the CFAF categories as ‘watercolors,’ and that great painters use every

color on their palette to varying degrees.

Suggestions For Future Research

Few investigators have studied feedback specificity or feedback magnitude in choral rehearsals, and findings from this study suggest that further research is needed. While some trends in feedback research seem to be emerging, many characteristics of feedback need to be empirically explored to establish a broader understanding of how different types of feedback are used in different settings. In particular, further study is needed to establish whether relationships exist between feedback characteristics and factors such as improved student attitude and performance, or rehearsal efficiency.

While several researchers have attempted to find relationships between teachers' use of complete sequential patterns on the aforementioned factors (Jellison & Kostka, 1987; Price, 1983, 1989), no study could be found on feedback specificity as it relates to student attitude and performance, or rehearsal efficiency. Researchers who use the CFAF may be able to provide data that identify certain feedback characteristics as more effective in ensemble rehearsals. Additionally, feedback data collected using the CFAF may be used to examine the effect specificity has on retention of concepts by students.

A future researcher might employ the use of an additional observation well after the primary investigation period to determine if participants continue to use specificity in their feedback. Though 96% of participants in the current study indicate that they continued to use specificity during subsequent rehearsals, the use of a confederate observer one month after the intervention period could provide more accurate empirical data. In addition, a longer observation period may reveal more accurate data representing participants' use of pre and post intervention specificity.

Future researchers may also examine feedback differences between choruses of different developmental levels, such as middle school and high school choral ensembles. While researchers have suggested that teachers give more approval to younger students (Kostka, 1984), those results have not been found consistently across various music classroom settings (Siebenaler, 1997, Speer, 1994). In addition, future researchers might use the CFAF to identify differences in feedback given by the same teacher to choral ensembles of varying developmental levels.

Finally, the results of this study seem to indicate the CFAF is an effective tool to both measure and affect change in conductor feedback behaviors. Though the CFAF was used exclusively with choral conductors in this study, the form could be adapted to analyze feedback tendencies of various types of teachers in a variety of settings. In addition, because no significant differences were observed between live-observation and distance-observation participants, the CFAF may also be applied in scenarios in which live observation is not possible, extending the reach of professional development. By examining teacher feedback in different musical environments, perhaps cross-curricular transfers can be made to increase instructional efficiency, improving student performance through feedback proficiency.

The present study suggests that: (a) specificity increases the magnitude of the feedback, and (b) the choral directors who participated tend to be highly general in their positive feedback to singers. If choral teachers can find ways to integrate specificity into their positive feedback, then they modify the function of feedback from simply informing the student that they did well, to informing the student *what* they did well. It is possible that by using more specificity in positive feedback, choral directors can simultaneously affect both the mood of rehearsals and the effectiveness of their instruction. • IJRCS

Institutional Review Board Approval and Compliance

The author obtained approval from an appropriate Institutional Review Board to conduct this research in a manner that assured the ethical treatment of participants and the confidentiality of participant information.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References

- Arnold, J. A. (1991). Using videotape self-analysis to improve teaching during rehearsals. *Update: Applications of Research in Music Education*, 10(1), 10-14.
- Arthur, J. R. (2002). *Experienced teachers' use of time in choral rehearsals of beginning and advanced choirs* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3109261)
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bennett P. D. (1989). Is praise always positive? *Connection (TMEC-MENC)*, 3(2), 12-13.
- Bergee, M. J. (1992). A scale assessing music student teachers' rehearsal effectiveness. *Journal of Research in Music Education*, 40(1), 5-13.
- Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, 51(1), 5-32.
- Byo, J. L. (1994). Comparison of expert and novice conductors: An approach to the analysis of nonverbal behaviors. *Journal of Band Research*, 30(1), 11-34.
- Carpenter, R. A. (1988). A descriptive analysis of relationships between verbal behaviors of teacher conductors and ratings of selected junior and senior high school band rehearsals. *Update: Applications of Research in Music Education*, 7(1), 37-40.
- Cavitt, M. E. (2003). A descriptive analysis of error correction in instrumental music rehearsals. *Journal of Research in Music Education*, 51(3), pp. 218-230.
- Derby, S. E. (2001). *Rehearsal of repertoire in elementary, middle, and high school choirs: How teachers effect change in student performance* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3008313)
- Duke, R. A., & Henninger, J. C. (1998). Effects of verbal corrections on student attitude and performance. *Journal of Research in Music Education*, 46(4), 482-495.
- Duke, R. A., & Henninger, J. C. (2002). Teachers' verbal corrections and observers' perceptions of teaching and learning. *Journal of Research in Music Education*, 50(1), 75-87.
- Duke, R. A., & Simmons, A. L. (2006). The nature of expertise; Narrative descriptions of 19 common elements observed in the lessons of three renowned artist-teachers. *Bulletin of the Council for Research in Music Education*, 170, 7-19.
- Forsythe, J. L. (1975). The effect of teacher approval, disapproval, and errors on student attentiveness: Music versus classroom teachers. In C. K. Madsen, R. D. Greer, & C. H. Madsen, Jr. (Ed.), *Research in music behavior* (pp. 49-55). New York: Teachers College Press.
- 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), 20-40.
- Grashel, J. (1994). Qualitative methodologies in music education research conference final session: Reflections and predictions. *Bulletin of the Council for Research in Music Education*, 122, 104.
- Hendel, C. (1995). Behavioral characteristics and instructional patterns of selected music teachers. *Journal of Research in Music Education*, 43(3), 182-203.
- Jellison, J. A., & Kostka, M. J. (1987, March). *Student or teacher as focus of attention and elementary students' written recall of teaching content within music teaching units*. Paper presented at the Research in Music Behavior Symposium, Logan, Utah.
- Kostka, M. J. (1984). An investigation of reinforcements, time use, and student attentiveness in piano lessons. *Journal of Research in Music Education*, 32(2), 113-122.
- Kuhn, T. L. (1975). The effect of teacher approval and disapproval on attentiveness, musical achievement, and attitude of fifth-grade students. In C. K. Madsen, R. D. Greer, & C. H. Madsen, Jr. (Ed.), *Research in music behavior* (pp. 40-48). New York: Teachers College Press.
- Langer, P. (1983, April). *Instruction and the feedback dilemma*. Unpublished paper presented at the meeting of the Rocky Mountain Psychological Association; Salt Lake City, UT.
- Madsen, C. K. (1990). Teacher intensity in relationship to music education. *Bulletin of the Council for Research in Music Education*, 104, 38-46.
- Madsen, C. K., & Alley, J. M. (1979). The effect of reinforcement on attentiveness: A comparison of behaviorally trained music therapists and other

- professionals with implications for competency-based academic preparation. *Journal of Music Therapy*, 16, 70-82.
- Madsen, C. K., & Duke, R. A. (1985). Observation of Approval/Disapproval in music: Perception versus actual classroom events. *Journal of Research in Music Education*, 33(3), 205-214.
- Madsen, C. K., & Duke, R. A. (Fall 1993). Selection and development of prospective music teachers. *Journal of Music Teacher Education*, 3(1), 5-11.
- Madsen, C. H., Jr., & Madsen, C. K. (1998). *Teaching/Discipline: Behavioral principles toward a positive approach*. Raleigh, NC: Contemporary Publishing.
- Morgan, R. (1992). *A study of a director's behaviors and his students' perceptions in a high school choral ensemble* (Doctoral dissertation). *Dissertation Abstracts International*, 53, 1836.
- Murray, K. C. (1975). The effect of teacher approval/disapproval on musical performance, attentiveness, and attitude of high school choruses. In C. K. Madsen, R. D. Greer, & C. H. Madsen, Jr. (Eds.), *Research in music behavior* (pp. 165-180). New York, NY: Teachers College Press.
- Napoles, J. (2006). The relationship between type of teacher talk and student attentiveness. *Journal of Music Teacher Education*, 16(1), 7-19.
- Price, H. E. (1983). The effect of conductor academic task presentation, conductor reinforcement, and ensemble practice on performers' musical achievement, attentiveness, and attitude. *Journal of Research in Music Education*, 31(4), 245-257.
- 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 verbal mannerism of song leaders. In C. K. Madsen & C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 125-134). Tuscaloosa, AL: University of Alabama Press.
- Siebenaler, D. J. (1997). Analysis of teacher-student interactions in the piano lessons of children and adults. *Journal of Research in Music Education*, 45(6), 6-20.
- Speer, D. R. (1994). An analysis of sequential patterns of instruction in piano lessons. *Journal of Research in Music Education*, 42(1), pp. 14-26.
- Whitaker, J. A. (2011). High school band students' and directors' perceptions of verbal and nonverbal teaching behaviors. *Journal of research in music education*, 59(3), 290-309.
- Yarbrough, C. (1975). Effect of magnitude of conductor behavior on students in selected mixed choruses. *Journal of Research in Music Education*, 23(2), 134-146.
- Yarbrough, C., & Hendel, C. (1993). The effect of sequential patterns on rehearsal evaluations of high school and elementary students. *Journal of Research in Music Education*, 41(3), 246-257.
- Yarbrough, C., & Madsen, K. (1998). The evaluation of teaching in choral rehearsals. *Journal of Research in Music Education*, 46(4), 469-481.
- Yarbrough, C., & Price, H. E. (1981). Prediction of performer attentiveness based on rehearsal activity and teacher behavior. *Journal of Research in Music Education*, 29(3), 209-217.
- Yarbrough, C., & Price, H. E. (1989). Sequential patterns of instruction in music. *Journal of Research in Music Education*, 37(3), 179-187.
- Yarbrough, C., Price, H. E., & Bowers, J. (1991). The effect of knowledge of research on rehearsal skills and teaching values of experienced teachers. *Update: Applications of Research in Music Education*, 9(2), 17-20.

This article is based on the author's Ph.D. dissertation, *Increasing Feedback Specificity in the Choral Rehearsal: Assessing and Enhancing Feedback of Teacher Conductors Through Analysis and Training*, defended in 2012 at The Florida State University.



Tucker Biddlecombe is an associate professor of music and Director of Choral Activities in the Blair School of Music at Vanderbilt University (USA). His research interests include conductor feedback, and voice usage in choral ensembles.