

## The Influences of Teacher Delivery and Student Progress on Preservice Teachers' Perceptions of Teaching Effectiveness.

By: Jessica Napoles and Rebecca B. MacLeod

Napoles, J. & [MacLeod, R. B.](#) (2013). The Influences of Teacher Delivery and Student Progress on Preservice Teachers' Perceptions of Teaching Effectiveness. *Journal of Research in Music Education*, 61(3): 249-261.

Made available courtesy of SAGE Publications:

<http://jrm.sagepub.com/content/61/3/249.abstract>

**\*\*\*Reprinted with permission. No further reproduction is authorized without written permission from SAGE Publications. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. \*\*\***

### **Abstract:**

The purpose of this study was to examine how teacher delivery and student progress influenced preservice teachers' perceptions of overall teaching effectiveness. Experienced teachers ( $n = 6$ ) were videotaped teaching mini applied lessons under four conditions: (a) high teacher delivery and more student progress, (b) high teacher delivery and less student progress, (c) low teacher delivery and more student progress, and (d) low teacher delivery and less student progress. Preservice teachers ( $n = 75$ ) viewed these teaching excerpts and rated each for teacher delivery, student progress, student musicianship, teacher knowledge of subject matter, and overall teaching effectiveness. Participants rated teachers with high delivery as more effective than teachers with low delivery, irrespective of student progress. There was a moderate positive correlation ( $r = .53$ ) between perceptions of teacher delivery and student progress. Results of a multiple regression analysis revealed that teacher delivery was the best predictor of perceptions of overall teaching effectiveness, followed closely by student progress.

**Keywords:** teacher delivery | student progress | teacher effectiveness | music education

### **Article:**

The assessment of teacher effectiveness recently has become the center of many conversations and debates related to public school teaching. There has been a strong movement in education towards teacher evaluations that focus on student performance—specifically, progress on standardized tests. In fact, the Marzano Evaluation Model, designed to establish a direct causal link between teaching and student achievement, is currently being employed by the Florida Department of Education as a model that districts can use or adapt to evaluate teachers (Marzano, 2003, 2006, 2007). The majority of testing and evaluation has been conducted relative to the “academic core,” and until recently, music and the arts have not been evaluated in the

same manner. However, the practice of evaluating teachers based on student progress is now becoming universal.

Research on music teacher effectiveness primarily has been focused on teacher behaviors and delivery of instruction, while few researchers have explored how instruction impacts student learning. Additionally, little research has been conducted to investigate whether student progress is an effective predictor of effective teaching or how student progress impacts viewers' perception of teaching effectiveness. Learning frequently takes place inside and outside the classroom, and numerous variables impact student performance on tests. Given recent No Child Left Behind policies (NCLB, 2002), government calls for increased accountability in the classroom as measured by student progress, and the controversial debate regarding teacher evaluation, it seems prudent to investigate the interaction between instruction and student outcomes.

Teacher delivery, teacher intensity, and conductor magnitude are terms that have been used to describe how the music teacher dispenses information. When defining teacher delivery, Hamann, Baker, McAllister, and Bauer (2000) included teacher posture, eye contact, gestures, facial expression, and vocal inflection. Yarbrough's (1975) definition of conductor magnitude was similar: eye contact, voice volume and modulation, proximity, facial expression, and pacing. In each case, participants preferred (rated higher in attitude) the teacher/conductor with good teacher delivery skills to the one with lower delivery skills. In fact, Hamann et al. found that participants preferred lessons with good teacher delivery, irrespective of lesson content. They concluded that university music students weighed quality of teacher delivery more heavily than the content of lesson. This finding was supported by K. Madsen (2003) with secondary students, undergraduate musicians, and experienced teachers when they evaluated elementary music lessons. All groups attended to delivery skills more than any other factor, and secondary students were influenced more by delivery than they were by the accuracy of instruction. Redding (2011) conducted a replication and extension of Madsen's study in a high school choral setting and found similar results. He included experienced nonmusic teachers in his sample, and they were most concerned with teacher delivery, perhaps because they could not tell when inaccurate information was being presented—a finding with strong implications for teacher evaluations conducted by administrators out of field. MacLeod and Napoles (2012) conducted a multiple regression study of factors predicting perceived teacher effectiveness, and teacher delivery was the best predictor among the variables investigated.

Several nonmusic studies have provided evidence that students in educational contexts were swayed by teacher delivery when evaluating teacher effectiveness. Burns, McKinney, and Burts (1985) examined preschoolers' attitudes and found that they were better toward social studies lessons with a "high-enthusiasm" teacher, compared to the low- and medium-enthusiasm teachers. In an older study, Coats, Swierenga and Wickert (1972) factor-analyzed student perceptions of educators in 1,427 7th- through 12th-grade classes. The researchers concluded that students regarded "teacher charisma" or "popularity" as the most important characteristic

when rating teachers. Delivery has a strong impact on teaching perceptions for music and nonmusic students alike.

Teacher intensity has been operationally defined as the sustained control of the student–teacher interaction, with efficient, accurate presentation of subject matter combined with enthusiastic affect and pacing (C. Madsen & Geringer, 1989). Hence, teacher intensity is different from conductor magnitude and teacher delivery in that it also encompasses classroom management skills and subject matter competence of the teacher (C. Madsen, 1990). Researchers studying teacher intensity have found that it was a teaching skill that could be taught, learned, and measured (Standley & Madsen, 1987), and it could be observed and rated reliably in others by participants ranging from middle school to college age (Byo, 1990; C. Madsen, Standley, & Cassidy, 1989). Training enhanced music education students' ability to maintain intensity while teaching music (Cassidy & Madsen, 1987), and there was a strong positive relationship between teacher intensity and teacher effectiveness ratings (C. Madsen, 1988). Lack of subject matter expertise contributed to less intensity (Cassidy, 1990), consistent with the inclusion of this component in the term's definition.

There has been a limited amount of research in music concerning student progress. Duke (1999/2000) noted in a review of 25 years of music education research (1972– 1997) that only 13 of 86 investigations measured student achievement. Others have postulated that teachers' ability to attend to student behavior rather than their own teaching behaviors appeared to increase with teaching experience (Bergee, 2005; Paul, 1998). While some have attempted to measure student progress (Siebenaler, 1997), or have students measure their own progress (Morrison, Montemayor, & Wiltshire, 2004), the association between teacher behavior and student achievement has not been established strongly (Duke, 1999/2000; Montemayor, 2006, 2011). Dickinson (1990) conducted an interesting study of student teachers teaching high school students in nonmusic subjects and found small correlations between instructor ratings and actual content learned; however, strong correlations were found between instructor ratings and perceived amount learned. Yarbrough and Henley (1999) asked an important question: "Is it possible for a teacher to get a good teaching evaluation if the performance of the students is poor?" (p. 309).

Several music researchers have endeavored to create tools for measuring student progress. Duke (1994) established the rehearsal frame as a means of observing how students achieved teacher-identified instructional goals. A rehearsal frame began "when the conductor first identified a problem in need of correction in the ensemble" and ended "when the identified problem was performed in its original context by the full ensemble" (Duke, 1994, p. 84). Montemayor (2006) formulated "rehearsal dyads," which were essentially snapshots before and after a musical goal was accomplished. He measured student progress using a Rehearsal Improvement Index score.

Although previous studies have indicated that teacher delivery has a strong influence on participants' perceptions, student progress has not been examined as a factor affecting

perceptions of effective teaching. Because the current assessment climate in the public schools is focused primarily on individual accountability and gains for each student, it seems important that preservice teachers are able to perceive student progress when it occurs. The primary purpose of this study was to determine how teacher delivery (high/low) and student progress (less/more) influence preservice teachers' perceptions of teaching effectiveness in private-lesson contexts. The private-lesson setting was used because it eliminated a number of variables and allowed preservice teachers to focus on one student. Specifically, our research questions were as follows: (1) Do preservice music teachers perceive teaching excerpts that contain high teacher delivery differently than excerpts containing low teacher delivery? (2) Do preservice teachers perceive teaching excerpts that contain more or less student progress differently? (3) Is there a difference in perceived teaching effectiveness between lessons with high teacher delivery compared to lessons with low teacher delivery? (4) Is there a difference in perceived teaching effectiveness between lessons with high student progress and low student progress? and (5) Is there a relationship between perceptions of teacher delivery and perceptions of student progress? A secondary purpose of this study was to determine which of the following variables influenced participants' perceptions of overall teaching effectiveness: teacher delivery, student musicianship, student progress, and teacher knowledge of subject matter.

## Method

### Participants

Participants were 75 music education majors from two large state universities, one in Utah and the other in North Carolina. Both universities were located in urban settings. Participants were enrolled in the current semester's offerings of music education courses. There were 26 females and 49 males, 12 freshmen, 30 sophomores, 25 juniors, and 8 seniors. Of the 75 participants, 12 identified themselves as choral in emphasis, and 63 identified themselves as instrumental.

### Preparing the Stimulus Recording

Six experienced teachers (brass,  $n = 2$ ; voice,  $n = 2$ ; and strings,  $n = 2$ ) were recruited to teach mini applied lessons to a beginning student (one of the researchers served as a mock student for all lessons). The teachers were between the ages of 28 and 45 and were Caucasian. Three teachers were female, and three were male. For purposes of consistency, all teachers were asked to teach their lesson and incorporate the following: (a) Announce two learning objectives at the beginning of the lesson, (b) include some type of teacher modeling on their instrument or voice, and (c) incorporate feedback that was contingent and specific, in a ratio of four approvals to one disapproval, which had been found to influence positively perceptions of teaching effectiveness in previous research (MacLeod & Napoles, 2012). Teaching excerpts were approximately 2 min in duration. There is evidence to suggest that participants can make adequate assessments of teaching with high agreement in a short amount of time through observing teachers' verbal (Mayhew, 2011) and nonverbal behaviors (Ambady & Rosenthal, 1993).

Teacher delivery and student progress were the two variables manipulated. Therefore, we created four conditions: high teacher delivery and more student progress, high teacher delivery and less student progress, low teacher delivery and more student progress, and low teacher delivery and less student progress. High teacher delivery was defined according to the behaviors identified in the research literature: eye contact, vocal modulation and volume, engaging facial expression, and energy. Given that each teacher announced two learning objectives for each lesson, student progress was defined operationally as the student's having accomplished both goals (more student progress) or having partially accomplished one of the two goals (less student progress).

Teachers were asked to dress professionally for their video recording, as the camera would be focused exclusively on them. A digital video camera (Canon FS400) recorded the teachers' lessons in a university faculty studio, with a view of the teachers' upper body only (head to torso). All (except vocalists) brought musical instruments and provided one for the student to use. In order to provide consistency among the lessons, we asked each teacher to identify and work on two goals for a beginning student. For strings, the two goals were boxed fingers and clear pizzicato tone. For vocalists, goals were tall vowels and singing with a supported tone. Brass players were to perform a sustained buzz on the mouthpiece and make a sound on the horn. Each teacher had several "takes" for each condition, and these were recorded for later review by independent observers, who were asked to choose a single excerpt that best reflected high/low teacher delivery and more/less student progress while exhibiting all stipulations listed earlier. Trials not meeting criteria were discarded.

#### Validity Check/Implementation of the Independent Variable

In order to determine whether the appropriate teacher delivery (high/low) and student progress (more/less) conditions had been executed, five experienced teachers, unfamiliar with the research study and the teachers in the video, were invited to serve as independent judges. Using Montemayor's (2006) dyad concept, the researchers created a before-and-after audio recording for each student. Judges listened to paired comparisons of the first student trial and final student trial from the mini teaching excerpts and were asked to determine whether more or less progress had been displayed by circling the appropriate response. The judges were not given detailed instructions regarding the performance goals that were established by the teacher in the video. Rather, they rated the student's progress according to their own paradigm. Similarly, the experts viewed two excerpts of the same teacher (one of high delivery and one of low delivery) in a paired comparison format and were asked to select the version they felt displayed high delivery.

Reliability among the five observers (computed using the formula  $[\text{agreements}]/[\text{total observations}]$ ) was 98% for teacher delivery. Delivery was rated as we had intended by all five judges in every instance but one, where one of the five judges disagreed. The reason for the one disagreement was unclear and may be attributed to fatigue, confusion, or personal preference. This high level of agreement indicated that the teacher delivery variable was implemented correctly and could be perceived as intended.

Reliability for student progress was 79%. There were four excerpts where judges were split in their assessments (3/2); thus, student progress was not perceived reliably as “less progress” as had been intended in these four instances. Therefore, we altered the four excerpts by replacing audio of the final student performance trial with an earlier performance trial from that particular lesson that demonstrated less progress. Because the student was hidden from view of the camera, the audio adjustments were not noticeable.

### Contents of Video Recording

After validity checks were complete and final selections were made, there were four excerpts for each of the six teachers: (a) high teacher delivery, more student progress (hereafter referred to as HD/MP); (b) high teacher delivery, less student progress (HD/ LP); (c) low teacher delivery, more student progress (LD/MP); and (d) low teacher delivery, less student progress (LD/LP). There were 24 excerpts and one practice example total. We created four presentation DVDs, two versions that contained 12 excerpts in each, wherein each disc included 6 high-teacher-delivery excerpts (three with more student progress, three with less) and 6 low-teacher-delivery excerpts (also three with more and three with less progress). Care was taken to place excerpts of each teacher’s high-teacher-delivery lessons on a different disc than the same teacher’s low-teacher-delivery lessons, to avoid the confusion that would occur if participants were to notice the extreme change in delivery style. The 12 excerpts contained in DVD 2 were in reverse order of excerpts on DVD 1, and DVD 4 was an exact reverse order of DVD 3, to control for probable learning effects. Each disc was 21 min in duration, including 24-s transitions between excerpts.

### Procedure

Participants viewed one of the four discs and were given the following instructions: “You will be viewing 12 excerpts of music lessons taught to beginning students. Please answer the 5 questions for each teaching excerpt.” Participants were asked to rate teacher delivery, student progress, student musicianship level, teacher knowledge of subject matter, and overall teaching effectiveness on 5-point Likert-type scales, with the anchors very low above the first number and very high above the last number. Following a practice example designed to accustom them to the task, participants were asked if they had any questions before beginning the actual study. Having two questions related to teacher behaviors and two related to student behaviors was deliberate, to allow participants to consider both teacher and student behaviors (as defined by the participants) when evaluating teaching effectiveness. Purposely, teaching effectiveness was not defined for the participants, so that their responses to the individual categories would yield a more accurate estimation of the elements they believed contributed to teaching effectiveness.

### Results

In preparation for data analysis, ratings for each teaching element were combined by delivery and progress conditions across teachers so that comparisons could be made between the four distinct combinations of delivery and progress: HD/MP, HD/LP, LD/MP, and LD/LP. Individual

participant ratings then were added together between like categories; for example, the teachers' delivery scores in the HD/MP category were combined into one score, the student progress scores in the LD/LP category became one score, and so on. A scale of 3 to 15 (1 to 5 for each of three teachers per order) was used, and there were three variables (teacher delivery, student progress, and overall teaching effectiveness) considered for each of the categories. The other two rating scales (student musicianship and teacher knowledge) were not used in the analysis. Because previous research (MacLeod & Napoles, 2012) indicated that instrument type (of the participant and of the teacher) did not affect evaluations of teaching, we did not use instrument as an independent variable.

In order to answer our first four research questions, we conducted a MANOVA with the four delivery and progress conditions as within-subjects factors and participant ratings for delivery, progress, and overall effectiveness as the three variates. Correlations among the three variates were moderately positive. Results indicated a significant main effect overall,  $F(9, 648) = 48.90$ ,  $p < .001$ , partial  $\eta^2 = .40$ . Follow-up univariate ANOVAs revealed significant main effects for delivery,  $F(3, 216) = 146.08$ ,  $p < .001$ , partial  $\eta^2 = .67$ ; progress,  $F(3, 216) = 122.91$ ,  $p < .001$ , partial  $\eta^2 = .63$ ; and overall effectiveness ratings,  $F(3, 216) = 136.02$ ,  $p < .001$ , partial  $\eta^2 = .65$ . Using Bonferroni adjustments for multiple comparisons (alpha level adjusted to .0125), every category's ratings were found to be significantly different from every other category's ratings. Analysis of means revealed that HD/MP teachers were rated highest in delivery ( $M = 12.74$ ,  $SD = 1.57$ ), followed by HD/LP teachers ( $M = 11.83$ ,  $SD = 1.70$ ), then LD/MP teachers ( $M = 8.89$ ,  $SD = 2.20$ ). LD/LP teachers were rated lowest ( $M = 8.19$ ,  $SD = 2.21$ ). High-delivery teachers were recognized clearly as exhibiting higher delivery than their low-teacher-delivery counterparts. Similarly, participants were able to discern between more and less student progress. Excerpts with HD/MP were rated highest in student progress ( $M = 12.26$ ,  $SD = 1.59$ ), followed by LD/MP ( $M = 10.98$ ,  $SD = 1.71$ ), then HD/LP ( $M = 10.05$ ,  $SD = 1.95$ ), and LD/LP rated lowest ( $M = 7.92$ ,  $SD = 2.12$ ). Participants rated progress higher in the "more-progress" excerpts than in the "less-progress" excerpts.

Participants considered HD/MP excerpts to be most effective ( $M = 12.63$ ,  $SD = 1.63$ ) and LD/LP excerpts to be least effective ( $M = 8.26$ ,  $SD = 2.06$ ). HD/LP excerpts ( $M = 10.95$ ,  $SD = 1.66$ ) and LD/MP excerpts ( $M = 10.23$ ,  $SD = 1.72$ ) were perceived similarly. Figure 1 (available online at <http://jrme.sagepub.com/supplemental>) provides a visual representation of comparisons between all categories for all three variables considered.

In order to answer our fourth research question, whether perceptions of teacher delivery were associated with perceptions of student progress, we conducted a Pearson correlation and found a moderate positive relationship between the two variables,  $r = .53$ ,  $p < .01$ . Mean ratings showed that participants rated student progress higher when teacher delivery was high compared to when teacher delivery was low, despite the fact that the amount of student progress was controlled and nearly identical in these examples (HD/LP,  $M = 10.05$ ,  $SD = 1.95$ ; and LD/LP,  $M = 7.92$ ,  $SD = 2.12$ ).

In an attempt to answer our final research question, which of the variables influenced perceptions of overall teaching effectiveness, we conducted a multiple regression analysis. Examinations of histograms and scatter plots indicated that the assumptions for linearity, normality, and homoscedasticity were met. Tolerance for each independent variable was above .1, indicating that multicollinearity assumptions also were met. All variables were entered simultaneously employing an exploratory regression model. An alpha level of .01 was established a priori, given the large data set.

Regression results indicated that the overall model significantly predicted teaching effectiveness,  $R^2 = .81$ , adjusted  $R^2 = .81$ ,  $F(4, 884) = 1001.26$ ,  $p < .001$ . This model accounted for 81.9% of variance in perceptions of overall teaching effectiveness. There were positive correlations (.32 to .79) between each predictor and overall teaching effectiveness. Bivariate and partial correlation coefficients are presented in Table 1 and illustrate that all four independent variables (teacher delivery, knowledge of subject matter, student progress, and student musicianship) had  $t$  values that significantly contributed to the model. Teacher delivery had the highest beta weight (.47), and the highest partial  $r$  (.63), closely followed by student progress ( $\beta = .42$ , partial  $r = .52$ ).

**Table 1.** Regression Coefficients for Final Model.

Variable	$B$	$\beta$	$t$	$p$	Bivariate $r$	Partial $r$
Teacher delivery	.40	.47	24.33	<.001	.79	.63
Student progress	.41	.42	19.41	<.001	.77	.54
Subject knowledge	.14	.12	6.39	<.001	.64	.20
Student musicianship	.06	.06	3.14	<.01	.53	.10

## Discussion

The primary purpose of this study was to examine how teacher delivery and student progress influenced preservice teachers' perceptions of overall teaching effectiveness. Results indicated that participants perceived teaching excerpts with high teacher delivery to be more effective than segments with low teacher delivery, irrespective of the progress condition. Excerpts with high teacher delivery and less student progress were rated higher in effectiveness than excerpts with low teacher delivery and more student progress. In answer to Yarbrough and Henley's (1999) question, based on the preservice teachers' ratings in this study, it appears that a teacher indeed is able to get a good teaching evaluation even when the student's performance is relatively poor. Additional research is needed to examine whether experienced teachers would respond similarly, or if they would be more sensitive to student progress behaviors, as suggested by Bergee (2005) and Paul (1998).



In addition to overall teaching effectiveness, participants rated the amount of student progress that was perceived in each video excerpt. Overall, teaching excerpts that contained more student progress were rated higher than excerpts that contained less student progress. However, when comparing HD/LP and LD/LP excerpts, participants rated progress much lower in the LD/LP excerpts ( $M = 15.40$ ) than in the HD/LP excerpts ( $M = 19.81$ ), despite the fact that student progress in those two examples was nearly identical. Similarly, participants rated delivery higher during the more-progress excerpts than during the less-progress excerpts (see Figure 1, at <http://jrme.sagepub.com/supplemental>). Delivery and progress appear to be intertwined, such that when one element is considered lacking, it affects viewers' perceptions of the other element, and vice versa. This interrelationship is corroborated further by the positive correlation between the two variables ( $r = .53$ ). Future research is needed to explore the impact of delivery on viewers' perception of student progress and whether experienced teachers would be influenced similarly by teacher delivery.

Results from the regression analysis revealed that teacher delivery was the best predictor of perceived overall teaching effectiveness, closely followed by student progress. Other researchers have noted the importance of teacher delivery in affecting participants' perceptions (Hamann et al., 2000; MacLeod & Napoles, 2012; C. Madsen, 1988; K. Madsen, 2003), so this finding was not too surprising. However, it is encouraging that student progress also had a large contribution and that it was perceived differentially, suggesting that delivery and progress were attended to distinctly by the preservice teachers. Although student musicianship significantly contributed to the overall model, the small beta weight and partial  $r$  lend credence to the fact that this variable was not nearly as important as the others. Teacher knowledge of subject matter also was perceived as important.

There were several limitations to this study. The excerpts were short in duration (1 to 2 min), and this may have affected participants' perceptions of what feasibly could be accomplished in the truncated lesson. Further, given that the camera was focused exclusively on the teachers, it is possible that overall teaching effectiveness was evaluated based on what the participants saw (teacher behaviors) rather than on what they heard (student behaviors). This limitation is especially important when considering that participants were rating string students' progress for displaying boxed fingers. Without the ability to see the students' fingers, accomplishment of this goal was difficult to evaluate. Authors of future studies may wish to incorporate a view of both the student and the teacher. Adding qualitative data, perhaps in the form of an open-ended question examining what participants were observing when rating overall teaching effectiveness, may elucidate further the elements to which they were attending.

Future research should include a comparison of experienced and preservice teachers' perceptions of student progress, teacher delivery, and overall teaching effectiveness. It is possible that the preservice teachers in this study were influenced by delivery in part due to their inexperience. Experienced teachers may attend more to student progress and may not be as influenced by teacher delivery. It also may be interesting to examine perceptions of nonmusic administrators,

because most public school teachers are evaluated by administrators unfamiliar with their content area. In addition, although we tried very carefully to control for content of lesson through consistency in teacher delivery and student progress, teacher effect was still evidenced. Given that 82% of the variance was accounted for in our regression model, it is possible that the remaining variance was attributable to participants' preference for certain teachers' presentation styles, personalities, and/or something else. Caution should be taken interpreting the results of this study and generalizing outside the context of preservice teachers' perceptions.

Results from this study carry important implications for music educators, music teacher evaluation, and music teacher educators. If teacher delivery influences perceptions of effectiveness so strongly as indicated in this study, then those presentation skills are crucial and should continue to be incorporated into methods classes. While one should not advise teachers to be less than competent in their subject matter, previous studies (MacLeod & Napoles, 2012; K. Madsen, 2003) have documented that knowledge of subject was less important than teacher delivery in evaluations of teachers.

Teacher evaluation in the United States currently is focused on student measures of achievement. Thus, it seems important that preservice teachers are able to recognize accurately when student progress occurs. In this study, the lesson excerpts evaluated were applied lessons that permitted the preservice teacher participants to attend to the progress of only one student at any given time. Yet, our findings indicated that preservice teachers' ratings of student progress appeared to be influenced by teacher delivery. More student progress was perceived in teaching excerpts that contained higher teacher delivery than in teaching excerpts that contained lower teacher delivery, despite the fact that the actual student progress in these excerpts was identical. Accurate observations of student progress seem important for teachers when delivering feedback and designing effective instruction. Therefore, it seems that preservice and experienced music teachers would benefit from developing and implementing assessment strategies that result in accurate measures of student progress.

Music teacher evaluation needs to be considered carefully to include accurate measures of teaching effectiveness. Especially because nonmusician administrators conduct much of this evaluation, it is important to continue to attend to the influences of teacher delivery and student progress and their interrelatedness. It is entirely possible that, if performed by individuals without training in the content area, evaluations will continue to be based primarily on teacher delivery. In light of this study's findings, high teacher delivery could lead evaluators to believe that students are progressing, even without actual progress occurring—a dangerous conclusion, at best. Clearly, evaluations of teacher effectiveness need to extend beyond a single measure. There appears to be a discrepancy between what policy makers value and what teacher educators have researched. While policy makers have focused on student progress to assess teacher effectiveness, researchers have focused on teacher delivery. It seems that a combination of these variables would be more accurate in measuring teacher effectiveness. Additional research is needed to identify a well-rounded model for music teacher evaluation.

## Declaration of Conflicting Interests

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

## Funding

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

## References

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64, 431–441. doi:0022-3514/93
- Bergee, M. J. (2005). An exploratory comparison of novice, intermediate, and expert orchestral conductors. *International Journal of Music Education*, 23, 23–26. doi:10.1177/0255761405050928
- Burns, D. C., McKinney, C. W., & Burts, B. L. (1985). Effects of teacher enthusiasm on three and four-year old children's acquisition of four concepts. *Theory and Research in Social Education*, 13, 19–32.
- Byo, J. L. (1990). Recognition of intensity contrasts in the gestures of beginning conductors. *Journal of Research in Music Education*, 38, 157–163. doi:10.2307/3345179
- Cassidy, J. W. (1990). Effect of intensity training on preservice teachers' instruction accuracy and delivery effectiveness. *Journal of Research in Music Education*, 38, 164–174. doi:10.2307/3345180
- Cassidy, J. W., & Madsen, C. K. (1987, November). The effect of training in high versus low presentation styles on teacher intensity. Paper presented at the national convention of the National Association of Music Therapy, San Francisco, CA.
- Coats, W. D., Swierenga, L., & Wickert, J. (1972). Student perceptions of teachers: A factor analytic study. *Journal of Educational Research*, 65, 357–360. Retrieved from <http://www.jstor.org/stable/27536307>
- Dickinson, D. J. (1990). The relationship between ratings of teacher performance and student learning. *Contemporary Educational Psychology*, 15, 142–151. Retrieved from [http://dx.doi.org/10.1016/0361-476X\(90\)90013-Q](http://dx.doi.org/10.1016/0361-476X(90)90013-Q)
- Duke, R. A. (1994). Bringing the art of rehearsal into focus: The rehearsal frame as a model for prescriptive analysis of rehearsal conducting. *Journal of Band Research*, 30, 78–95.

Duke, R. A. (1999/2000). Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education*, 143, 1–48. Retrieved from <http://www.jstor.org/stable/40319011>

Hamann, D. L., Baker, D. S., McAllister, P.A., & Bauer, W. I. (2000). Factors affecting university music students' perceptions of lesson quality and teaching effectiveness. *Journal of Research in Music Education*, 48, 102–113. Retrieved from <http://www.jstor.org/stable/3345569>

MacLeod, R. B., & Napoles, J. (2012). Preservice teachers' perceptions of teaching effectiveness during teaching episodes with positive and negative feedback. *Journal of Music Teacher Education*, 22, 91–102. doi:10.1177/1057083711429851

Madsen, C. K. (1988). The relationship of teacher on-task to perceived teacher effectiveness. In P. J. Flowers (Ed.), *Proceedings of the Current Issues in Music Education Symposium*, Vol. 13 (pp. 2-8), Columbus, OH: School of Music, The Ohio State University.

Madsen, C. K. (1990). Teacher intensity in relationship to music education. *Bulletin of the Council for Research in Music Education*, 104, 38–46. Retrieved from <http://www.jstor.org/stable/40318356>

Madsen, C. K., & Geringer, J. M. (1989). The relationship of teacher “on task” to intensity and effective teaching. *Canadian Journal of Research in Music Education*, 30, 87–94.

Madsen, C. K., Standley, J. M., & Cassidy, J. W. (1989). Demonstration and recognition of high and low contrasts in teacher intensity. *Journal of Research in Music Education*, 37, 85–92. doi:10.2307/3344700

Madsen, K. (2003). The effect of accuracy of instruction, teacher delivery, and student attentiveness on musicians' evaluation of teacher effectiveness. *Journal of Research in Music Education*, 51, 38–50. Retrieved from <http://www.jstor.org/stable/3345647>

Marzano, R. J. (2003). *What works in schools*. Alexandria, VA: ASCD.

Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: ASCD.

Marzano, R. J. (2007). *The art and science of teaching*. Alexandria, VA: ASCD.

Mayhew, P. (2011). The effect of initial impressions on the perception of teaching effectiveness in choral music student teachers (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (Publication No. 3462329)

Montemayor, M. (2006). Rehearsal achievement in high school bands and its relationship to performance quality, selected rehearsal procedures, and evaluations of teaching effectiveness (Unpublished doctoral dissertation). University of Washington, Seattle.

Montemayor, M. (2011, February). Evaluation of intrarehearsal achievement by listeners of varying levels of expertise. Paper presented at the International Symposium for Research in Music Behavior, Barcelona, Spain.

Morrison, S. J., Montemayor, M., & Wiltshire, E. S. (2004). The effect of a recorded model on band students' performance self-evaluations, achievement, and attitude. *Journal of Research in Music Education*, 52, 116–129. doi:10.2307/3345434

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Paul, S. J. (1998). The effects of peer teaching experiences on the professional teacher role development of undergraduate instrumental music education majors. *Bulletin of the Council for Research in Music Education*, 137, 73–92. Retrieved from <http://www.jstor.org/stable/40318933>

Redding, J. B. (2011). The effect of choral conductor intensity on the perception of teacher effectiveness (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (Publication No. 3483596)

Siebenaler, D. J. (1997). Analysis of teacher-student interactions in the piano lessons of adults and children. *Journal of Research in Music Education*, 45, 6–20. doi:10.2307/3345462

Standley, J. M., & Madsen, C. K. (1987, Summer). Intensity as an attribute of effective therapist/client interaction. *Quodlibet*, pp. 15–20.

Yarbrough, C. (1975). Effect of magnitude of conductor behavior on students in selected mixed choruses. *Journal of Research in Music Education*, 23, 134–146. doi:10.2307/3345286

Yarbrough, C., & Henley, P. (1999). The effect of observation focus on evaluations of choral rehearsal excerpts. *Journal of Research in Music Education*, 47, 308–318. doi:10.2307/3345486