

The Influences of Teacher Delivery and Student Progress on Experienced Teachers' Perceptions of Teaching Effectiveness

By: [Rebecca B. MacLeod](#), Jessica Nápoles

MacLeod, R. B., & Napoles, J. (2014). The influence of teacher delivery and student progress on experienced teachers' perceptions of teaching effectiveness. *Journal of Music Teacher Education*. doi: 10.1177/1057083714527111

Made available courtesy of Sage Publications: <http://dx.doi.org/10.1177/1057083714527111>

*****Reprinted with permission. No further reproduction is authorized without written permission from Sage Publications & MENC: The National Association for Music Education and Society for Music Teacher Education. 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 replication and extension study was to examine how teacher delivery and student progress influenced experienced teachers' perceptions of overall teaching effectiveness. Participants ($N = 60$ experienced music teachers) viewed 12 private lesson excerpts that included four separate 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. Participants rated teacher delivery, student progress, student musicianship, teacher knowledge of subject matter, and overall teaching effectiveness for each private lesson excerpt. Teaching excerpts with high teacher delivery were rated as more effective than excerpts with low teacher delivery, irrespective of student progress. Results of a multiple regression indicated that teacher delivery was the largest predictor for experienced teachers' ratings of overall teaching effectiveness, followed closely by student progress.

Keywords: teacher effectiveness | experienced teachers | delivery | student progress | perceptions | music education

Article:

Teacher delivery has received a considerable amount of attention by music and non-music researchers, and investigations have dated back to the past 40 years or so. Many university teacher training curricula include heavy emphasis on teacher delivery skills as a measure of teaching effectiveness. However, teacher delivery is not the only variable to consider when evaluating overall effectiveness, nor is there strong evidence that teacher delivery has any impact on student learning.

The Dr. Fox studies conducted in the 1970s examined the issue of teacher delivery and its relationship to student evaluations of teacher effectiveness. In the original study, Naftulin, Ware, and Donnelly (1973) trained an actor (Dr. Fox) to deliver an articulate, entertaining, charismatic, and humorous lecture (on mathematical game theory as applied to physician education) without substance or accuracy in academic content. Participants (higher education professionals from various fields, none mathematics) rated the lecturer as highly effective. The authors concluded that style was more influential than content and that student satisfaction with learning may represent little more than the illusion of having learned. They labeled this phenomenon “educational seduction.”

In a follow-up study, Ware and Williams (1975) investigated the impact of content (high-, medium-, and low-content lectures) and levels of seduction (high-seduction lecturers employed enthusiasm, humor, friendliness, expressiveness, charisma, and personality, whereas low-seduction lecturers did not) on participants’ perceptions of teacher effectiveness. Similar to the previous Dr. Fox study, college students evaluated the lecturer’s knowledge and presentation manner at the end of the lesson. However, in this study the participants also completed a test over the lecture topic. This study’s findings indicated that, not only did students prefer the high-seduction lecturer, they also performed better on the cognitive test than students who viewed the low-seduction lecturer. The researchers asserted, “The Doctor Fox Effect appears to be more than an illusion. Seductiveness affects both student ratings of instruction and achievement” (Ware & Williams, 1975, p. 149). Subsequent studies (Ware & Williams, 1977; Williams & Ware, 1977) further expanded on and supported the Dr. Fox effect, but other researchers (Marsh & Ware, 1982; Meier & Feldhusen, 1979; Perry, Abrami, & Leventhal, 1979) found methodological flaws in the original studies and failed to replicate the results with real professors. Thus, although the “Dr. Fox effect” is debatable, there is evidence that students prefer teachers with high expressiveness to teachers with low expressiveness and that delivery is an important element of perceived teaching effectiveness. These findings could have potentially important implications for music teachers.

Similar to the Dr. Fox studies, music researchers have also explored the relationship between teacher delivery and student ratings of instruction. Teacher delivery has been defined through observable behaviors (posture, eye contact, gestures, facial expression, and vocal inflection) by numerous researchers, and terms such as *intensity* and *magnitude* have also been associated with similar behaviors (C. K. Madsen, 1988; C. K. Madsen, Standley, & Cassidy, 1989; Yarbrough, 1975). Like teacher delivery, teacher intensity has been defined according to enthusiastic affect and the student–teacher interaction but additionally included accurate presentation of subject matter and classroom management skills (C. K. Madsen, 1988). C. K. Madsen et al. (1989) found that intensity as a concept could be operationally defined, easily taught to prospective student teachers, ably demonstrated, and recognized with an extremely high degree of reliability. Yarbrough (1975) defined “magnitude” as including eye contact, closeness, volume and modulation of voice, gestures, facial expression, and rehearsal pace. Teacher delivery, teacher intensity, and conductor magnitude all appear to share common elements that affect perceptions of effective music teaching.

Through investigations focused on teacher delivery, teacher intensity, and conductor magnitude, researchers have found that not only is teacher delivery related to ratings of overall effectiveness,

delivery seems to take priority over accuracy and quality of content when evaluating the effectiveness of instruction. For example, Yarbrough and Madsen (1998) found that students may forgive inaccuracies in task presentations if the teacher has a satisfactory or pleasing style of teaching. In their observation, a college choral conductor modeled inaccurate rhythms in one segment, yet participants rated the teacher as effective. Hamann, Baker, McAllister, and Bauer (2000) also found that university students preferred teachers with high delivery skills, irrespective of lesson content. K. Madsen (2003) studied perceptions of secondary students, undergraduate musicians, and experienced teachers and found that all groups attended to delivery skills more than they did to accuracy of instruction. In fact, teaching excerpts with high delivery and inaccurate instruction were rated higher than excerpts with low teacher delivery and accurate instruction.

Teaching experience has been documented as an important factor affecting perceptions of teaching effectiveness. K. Madsen's (2003) secondary student participants rated excerpts with high teacher delivery, good classroom management, and *inaccurate* lesson content as very effective, higher than experienced teachers did. Redding (2011) replicated these results in a secondary choral rehearsal context, adding experienced teachers that were not music specialists, and they also responded differentially. Graduate students tended to rate interest and preference for teaching examples higher than did undergraduate students (Hamann et al., 2000). The researchers suggested that experience be isolated as a variable when evaluating teaching.

Student achievement has not been examined as closely (Duke 1999/2000; Montemayor, 2006), although instruments have been created for the purpose of measuring student progress in piano lessons (Siebenaler, 1997) and rehearsals (Bergee, 1992; Duke, 1994; Montemayor, 2006; Morrison, Montemayor, & Wiltshire, 2004). Napoles and MacLeod (2013) recently designed a study to examine how student progress and teacher delivery influenced perceptions of overall teaching effectiveness. Preservice teachers were able to differentiate between more and less progress, but the variable of teacher delivery affected their perceptions of overall teaching effectiveness more than student progress did. These preservice teachers rated lessons where a student demonstrated less progress and the teacher displayed high delivery as more effective overall than lessons with more progress and low teacher delivery.

The present study is a replication and extension of the Napoles and MacLeod (2014) study, with experienced music teachers as participants. The primary purpose of this study was to examine how teacher delivery (high/low) and student progress (more/less) influenced experienced teachers' perceptions of overall teaching effectiveness. The following research questions were included:

Research Question 1: Will experienced music teachers perceive excerpts with high teacher delivery differently than excerpts with low teacher delivery?

Research Question 2: Will experienced music teachers perceive excerpts with more student progress differently than excerpts with less student progress?

Research Question 3: Will experienced music teachers perceive a difference in teaching effectiveness between lessons with high teacher delivery compared to lessons with low teacher delivery?

Research Question 4: Will experienced music teachers perceive a difference in teaching effectiveness between lessons with more student progress and less student progress?

Research Question 5: Will there be a relationship between experienced music teachers' perceptions of teacher delivery and perceptions of student progress?

A secondary purpose, as with the original study, was to determine which of the following variables best predicted participants' perceptions of overall teaching effectiveness: teacher delivery, student musicianship, student progress, and teacher knowledge of subject matter. We added teaching experience as a predictor variable, to determine whether number of years teaching affected participant ratings.

Method

Participants

Participants were experienced music teachers recruited from students enrolled in the graduate degree programs at the University of Utah and the University of North Carolina (UNCG) at Greensboro as well as experienced teachers who were the instructors at the UNCG summer music camps. Of the original 65 teachers secured, 3 were dropped due to insufficient teaching experience (fewer than 3 years), another was dropped because he recognized one of the teachers in the video from another event, and a fifth teacher was dropped because of incomplete participation. Thus, there were 60 total participants, 19 males and 41 females, each with a minimum of 3 years teaching experience ($M = 9.51$, $Mdn = 7$, $SD = 6.45$), in choral ($n = 6$), general ($n = 4$), instrumental ($n = 19$), applied teaching ($n = 2$), or some combination thereof ($n = 29$).

The Stimulus Recording

We used videos from a previous study (Napoles & MacLeod, 2013) that included six experienced teachers in brass ($n = 2$), voice ($n = 2$), and strings ($n = 2$), teaching abbreviated applied lessons to beginning students. Half of the teachers were female, the other half male. All were Caucasian, and between the age of 28 and 45 years. Excerpts were approximately 2 minutes in duration. Teachers were asked to announce two learning goals at the beginning of the lesson, model on their instrument (or voice), and incorporate feedback that was specific and contingent, with four approvals and one disapproval in each lesson. Each teacher provided four lessons, one in which the teacher displayed high delivery and the student displayed more progress (accomplishing the two goals the teacher announced at the beginning), another in which the teacher displayed high delivery and the student displayed less progress (partially accomplishing one of the two goals), another in which the teacher displayed low delivery and the student displayed more progress, and another in which the teacher displayed low delivery and the student displayed less progress. These four conditions will subsequently be referred to as HD/MP (high delivery, more progress), HD/LP (high delivery, less progress), LD/MP (low delivery, more progress), and LD/LP (low delivery, less progress). High teacher delivery was operationally defined according to the behaviors reported by Hamann et al. (2000): expanded posture, frequent

eye contact, use of physical gestures, engaging facial expression, and changing voice volume and inflection.

There were 24 excerpts and one practice example total. We then created four presentation DVDs, two versions that contained 12 excerpts in each. The 12 excerpts contained on DVD 2 were in reverse order of the excerpts on DVD 1, and DVD 4 was an exact reverse order of DVD 3, to control for probable order effects. Each disc was approximately 21 minutes in duration, including 24-second transitions between excerpts.

Validity Check/Implementation of the Independent Variable

We performed a validity check when the videos were originally produced, with five experienced teachers serving as independent judges. Using Montemayor's (2006) dyad concept, we 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 abbreviated teaching excerpts and were asked to determine whether more or less progress had been displayed by circling the appropriate response. Similarly, the expert panel viewed two excerpts of the same teacher (one of high delivery and one of low delivery) in a paired comparison format and was asked to select the version they felt displayed high delivery.

Whereas reliability among the five observers—computed using the formula (agreements)/(total observations)—was 98% for teacher delivery and considered acceptable, reliability for student progress was 79%. There were four excerpts where judges were split in their assessments (3:2). Therefore, the researchers 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 students were hidden from view of the camera, the adjustments were not noticeable.

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. There were none. Having two questions related to teacher behaviors and two related to student behaviors was deliberate, to encourage 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 because our goal was to explore how ratings of delivery and student progress, regardless of individual value systems, might influence overall effectiveness ratings. Although all participants were asked to rate teacher delivery and student progress, it is possible that their own paradigm of teacher effectiveness included other variables.

Results

Before analyzing the data, we combined ratings across teachers for each of the four delivery and progress conditions: high delivery/more progress (HD/MP), high delivery/less progress (HD/LP), low delivery/more progress (LD/MP), and low delivery/less progress (LD/LP). We then added the individual participant ratings for delivery, progress, and overall effectiveness between like categories, so that the three teachers' delivery scores in the HD/MP category became one score, the three student progress scores in the LD/MP category became one score, etc. A scale of 3 to 15 was used—1 to 5 for each of the three data points—and there were three variables (teacher delivery, student progress, and overall teaching effectiveness) considered for each of the categories. The ratings for student musicianship and teacher knowledge were only relevant for the regression analysis.

Analysis of means revealed that HD/MP teachers were rated highest in delivery ($M = 12.60$, $SD = 1.54$), followed by HD/LP teachers ($M = 11.76$, $SD = 2.03$), then LD/MP teachers ($M = 8.01$, $SD = 2.28$). LD/LP teachers were rated lowest ($M = 7.36$, $SD = 2.60$). High delivery teachers were clearly recognized 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 ($M = 12.30$, $SD = 1.66$), followed by LD/MP ($M = 10.50$, $SD = 1.95$), then HD/LP ($M = 9.16$, $SD = 1.87$), and LD/LP rated lowest ($M = 7.36$, $SD = 2.39$). Participants rated progress higher in the “more progress” excerpts than in the “less progress” excerpts.

Participants considered HD/MP excerpts to be most effective ($M = 12.51$, $SD = 1.55$) and LD/LP excerpts to be least effective ($M = 7.71$, $SD = 2.83$). HD/LP excerpts ($M = 10.56$, $SD = 1.89$) and LD/MP excerpts ($M = 9.86$, $SD = 2.36$) were perceived similarly. Figure 1A provides a visual representation of comparisons between all categories for all three variables considered.

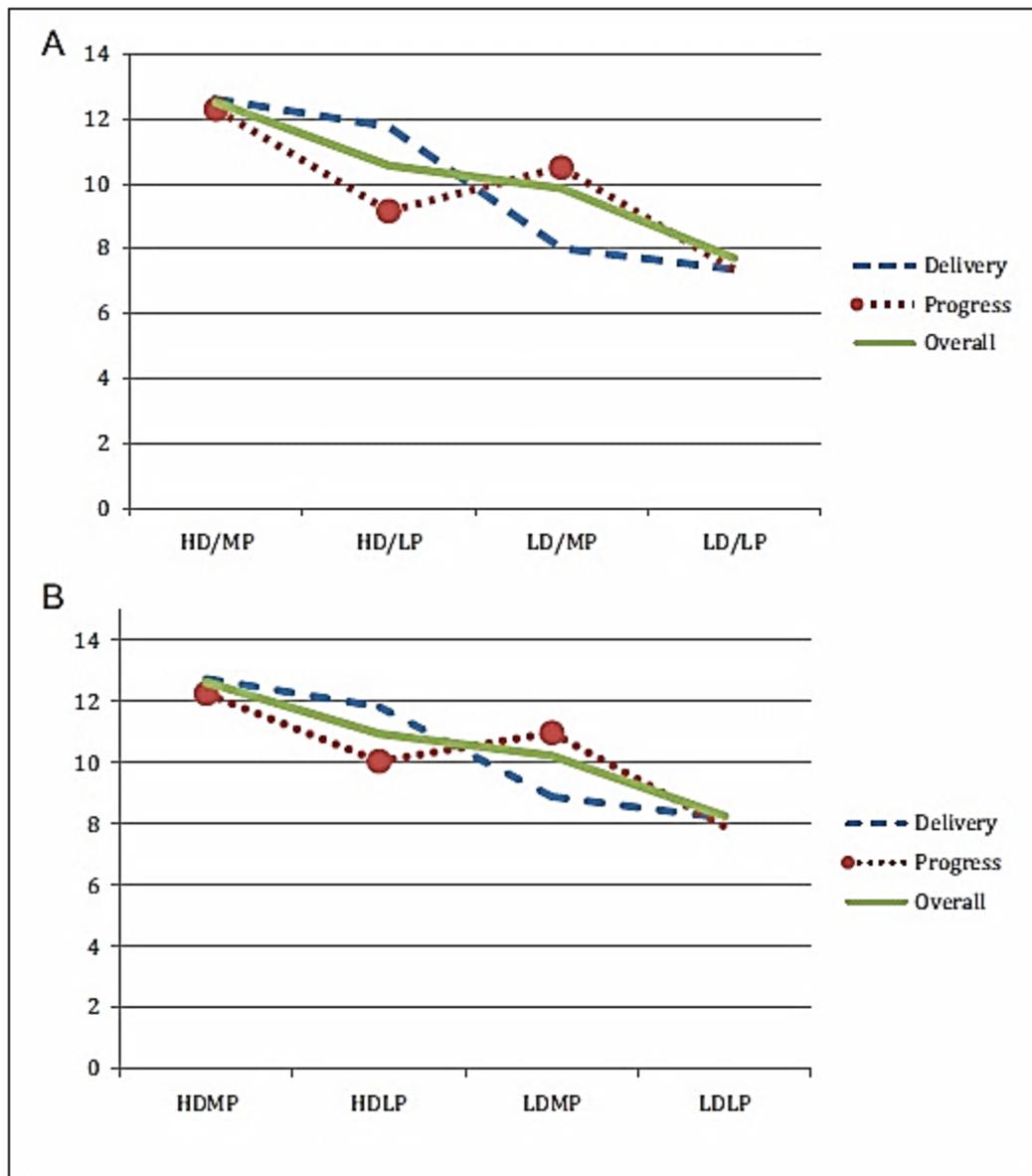


Figure 1. (A) Experienced teachers' and (B) preservice teachers' mean ratings for teacher delivery, student progress, and overall effectiveness. *Note.* HD = high delivery; MP = more progress; LD = low delivery; LP = less progress.

To answer our first four research questions, we conducted a multiple analysis of variance 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, Wilk's $\lambda = .09$, $F(9, 426) = 78.24$, $p < .001$, $\eta^2 = .91$. Follow-up univariate analyses of variance (ANOVAs) revealed significant main effects for delivery, $F(3, 177) = 188.80$, $p < .001$, $\eta^2 = .31$, progress, $F(3, 177) = 118.97$, $p < .001$, $\eta^2 = .20$, and overall effectiveness ratings, $F(3, 177) = 103.30$, $p <$

.001, $\eta^2 = .18$. Using Bonferroni adjustments for multiple comparisons, every category's ratings were found to be significantly different from every other category's ratings, $p < .01$, except delivery ratings between LD/MP and LD/LP excerpts, and overall effectiveness ratings between HD/LP and LD/MP excerpts.

To answer our fifth 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, $r = .53$, $p < .0001$. 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 in these examples was controlled, HD/LP, $M = 9.16$, $SD = 1.87$; LD/LP, $M = 7.36$, $SD = 2.39$.

To answer our final research question, which variables predicted perceptions of overall teaching effectiveness, we conducted a stepwise multiple regression. Scatter plots and histograms were examined for violations of linearity, homoscedasticity, and normality. Tolerance for each independent variable was above .1, indicating that multicollinearity assumptions were also met. No outliers were identified by calculating Mahalanobis distances. An alpha level of .01 was established a priori, given the large data set.

The model summary indicates that three of the five independent variables were entered into the model. For the first step, teacher delivery was entered as it accounted for the most unique variance in the independent variable ($R^2 = .64$). The variables of student progress and teacher knowledge were entered in the next two steps, respectively, creating a model that accounted for 83% of the variance in perceptions of overall teaching effectiveness. The ANOVA table (see Table 1) presents the F test for each step/model. The final model significantly predicts perceptions of overall teaching effectiveness, $F(3, 715) = 1151.98$, $p < .001$. A summary of the regression model is presented in Table 2. In addition, bivariate and partial correlation coefficients between each predictor and the dependent variable are presented in Table 3. Student musicianship and teaching experience did not contribute to the model.

Discussion

The purpose of this study was to examine how teacher delivery and student progress influenced experienced teachers' perceptions of overall teaching effectiveness. Experienced teachers rated the overall teacher effectiveness of the four conditions as follows: HD/MP excerpts were considered most effective ($M = 12.51$, $SD = 1.55$) and LD/LP excerpts least effective ($M = 7.71$, $SD = 2.83$). HD/LP excerpts ($M = 10.56$, $SD = 1.89$) and LD/MP excerpts ($M = 9.86$, $SD = 2.36$) were perceived similarly and not significantly different from one another. Similar to the findings of Napoles and MacLeod (2013) with preservice teachers (see Figure 1B), the experienced teachers in this study perceived teaching excerpts with high teacher delivery to be more effective than segments with low teacher delivery, irrespective of the progress condition.

Table 1. ANOVA Table for Regression.

	Sum of squares	df	Mean square	F	p
--	----------------	------	-------------	-----	-----

Model 1					
Regression	655.04	1	655.04	1282.78	<.001
Residual	366.13	717	.51		
Total	1021.17	718			
Model 2					
Regression	824.79	2	412.39	1503.53	<.001
Residual	19a6.38	716	.27		
Total	1021.17	718			
Model 3					
Regression	846.12	3	282.04	1151.98	<.001
Residual	175.05	715	.24		
Total	1021.17	718			

Note. ANOVA = analysis of variance; *df* = degrees of freedom.

Table 2. Regression Model Summary.

Step	<i>R</i>	<i>R</i> ²	<i>R</i> ² _{adj}	ΔR^2	<i>F</i> _{chg}	<i>p</i>	<i>df</i> ₁	<i>df</i> ₂
1. Teacher delivery	.80	.64	.64	.64	1282.78	<.001	1	717
2. Teacher delivery and student progress	.89	.80	.80	.16	618.86	<.001	1	716
3. Teacher delivery, student progress, and teacher knowledge	.91	.83	.83	.02	87.13	<.001	1	715

Note. *df* = degrees of freedom.

Table 3.

B	β	<i>t</i>	<i>p</i>	Bivariate <i>r</i>	Partial <i>r</i>	
Delivery	.40	.44	21.28	<.001	.80	.62
Progress	.45	.42	22.36	<.001	.76	.64
Knowledge	.23	.19	9.33	<.001	.71	.33

Furthermore, experienced teachers rated excerpts with high teacher delivery and less student progress higher in effectiveness than excerpts with low teacher delivery and more student progress. Some researchers have found evidence to support the idea that preservice teachers and experienced teachers observe and perceive music teaching and teacher effectiveness differently (Berger, 2005; Paul, 1998). However, the results of this study indicate that experienced teachers rated teacher effectiveness in the areas of teacher delivery, student progress, student musicianship, teacher knowledge of subject matter, and overall teaching effectiveness similarly to the preservice teachers in our previous study (Napoles & MacLeod, 2013).

Results of this study are similar to those of K. Madsen (2003) in that actual teacher delivery (as manipulated in the videos) strongly affected perceptions of overall teaching effectiveness. However, in her study, there was more differentiation between the perceptions of experienced teachers and preservice teachers. For example, if a teacher delivered inaccurate information, preservice teachers were more forgiving and rated the teacher as effective if the delivery was high. Experienced teachers, though, rated that same teacher as less effective, perhaps

understanding that there were other variables to consider in addition to teacher delivery. Experienced teachers in the current study did not evidence this degree of differentiation in perceptions, but it is important to note that the teaching excerpts in our study did not incorporate inaccurate instruction. Perhaps accurate content is more important to experienced teachers than student growth that is contained within a brief lesson excerpt. Regardless, the “Doctor Fox effect” was evidenced to some degree with these teachers in that delivery style appeared to influence teacher effectiveness ratings.

Although teacher delivery is an important element of effective teaching, we were particularly interested in experienced teachers’ ability to assess student progress and whether they might evaluate student progress more accurately than preservice teachers in our previous study. Overall, teaching excerpts that contained more student progress were rated higher than excerpts that contained less student progress. Indeed, participants were able to discriminate between the two conditions. However, when comparing HD/LP and LD/LP excerpts, participants rated *progress* lower in the LD/LP excerpts ($M = 7.36, SD = 2.39$) than in the HD/LP excerpts ($M = 9.16, SD = 1.87$), despite the fact that student progress in those two examples was nearly identical. The experienced teachers’ ratings were similar to the preservice teachers’ ratings of the same teaching excerpts ($M = 7.92$ for LD/LP excerpts; $M = 10.05$ for HD/LP excerpts) in our previous study. We found a moderately positive correlation between ratings of teacher delivery and student progress. It appears that both experienced teachers’ and preservice teachers’ ratings of student progress are related to ratings of teacher delivery. The interrelationship between these two variables is intriguing, when considering that behaviors like vocal inflection and eye contact are associated with perceptions of whether a student is achieving his or her music goals.

Results of the stepwise multiple regression revealed that three (teacher delivery, student progress, and teacher knowledge of subject matter) of the five independent variables were predictive of ratings for effective teaching. Student musicianship and teaching experience did not significantly predict the participants’ ratings of effectiveness. Teacher delivery was the strongest predictor of teacher effectiveness, accounting for 64% of the variance. Teacher delivery combined with student progress accounted for 80% of the variance, and when teacher knowledge of the subject matter was included in the model, these three variables accounted for 83% of the variance. Had additional variables been included, we may have found different results. Future research may include teacher musicianship, student engagement, or any number of variables that may affect perceptions of teaching effectiveness.

Conclusions and Recommendations

The current study was a replication and extension of a previous study that investigated preservice teachers’ perceptions of teacher effectiveness (Napoles & MacLeod, 2013). In the original study, teacher delivery was a strong predictor for preservice teachers’ perceptions of teacher effectiveness, closely followed by student progress. We were interested in investigating the perceptions of experienced teachers to determine whether the results of the first study were specific only to preservice teachers. Previous research indicated that preservice teachers rated the same teaching event differently than experienced teachers did (Bergee, 2005; Hamann et al., 2000; K. Madsen, 2003; Paul, 1998). Therefore, perhaps the most interesting finding in the current study was that experienced teachers perceived the private lessons very similarly to

preservice teachers. Close comparisons of mean ratings by preservice and experienced teachers for each of the three independent variables revealed very few differences (see Figure 1A and B). Both experienced and preservice teachers rated the excerpts with high teacher delivery as more effective than the excerpts with low teacher delivery. Also, both populations were able to discriminate between more and less student progress. Overall teacher effectiveness ratings were very similar. We also found the same correlation between teacher delivery and student progress ratings in both studies, indicating that these two variables were related for both sets of participants.

Based on the results of the current study, it seems that teacher delivery continues to be an extremely important element affecting perceptions of effective teaching. One may perhaps conjecture that preservice teachers would be more swayed by delivery given their inexperience, but in fact, this study's findings confirm that experienced teachers were swayed similarly. Teacher education programs should continue to focus on teacher delivery skills as they appear to affect how teaching effectiveness is perceived. However, it is equally important that teacher educators stress the importance of the *purpose* of instruction—student learning. It is not enough for teachers to display “high magnitude” or “teacher intensity” if the students do not make progress in learning.

The ability to assess a student's progress during a lesson is an essential skill for teachers, and this information should be used to design instruction accordingly. If observations of student progress in this study had been accurate, then student progress ratings would have been more similar in the HD/MP and LD/MP examples as the students in these examples made the same amount of progress. However, both preservice and experienced teachers rated the amount of progress in these examples differently, providing evidence that accurate assessment of student progress is not immediately and easily performed. A keen awareness of individual students' progress is extremely important to student success. For performance-based evaluation, it may be helpful for teachers to isolate a student's performance for assessment purposes so that attention can be exclusively focused on one student, as it is clear that outside variables may influence perceptions of the amount of progress a student is making. Additionally, teacher education programs should provide opportunities for preservice teachers to assess student learning both during and outside the context of teaching “in the moment” so that they may develop the skills necessary to accurately assess student progress.

Future research may examine evaluation of student progress more specifically, especially given the climate of tying teacher evaluations to student achievement. Regardless of participants' preferences for high teacher delivery, issues of student progress cannot and should not be ignored. Teacher educators/researchers must make earnest attempts find effective ways to measure student progress and to include student outcomes as dependent variables in empirical studies about teaching effectiveness.

Declaration of Conflicting Interests

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

Funding

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

References

Bergee, M. J. (1992). A scale assessing music student teachers' rehearsal effectiveness. *Journal of Research in Music Education, 40*, 5-13. doi:10.2307/3345770

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

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.

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. doi:10.2307/3345569

Madsen, C. K. (1988). The relationship of teacher on-task to perceived teacher effectiveness. *Proceedings of the Current Issues in Music Education Symposium, 13*, 2-8.

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.

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. doi:10.2307/3345647

Marsh, H. W., & Ware, J. E., Jr. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of the Dr. Fox effect. *Journal of Educational Psychology, 74*, 126-134.

Meier, R. S., & Feldhusen, J. F. (1979). Another look at Dr. Fox: Effect of stated purpose for evaluation, lecturer expressiveness, and density of lecture content on student ratings. *Journal of Educational Psychology, 71*, 339-345.

Montemayor, M. (2006). *Rehearsal achievement in high school bands and its relationship to performance quality, selected rehearsal procedures, and evaluations of teaching effectiveness*. Retrieved from Proquest Dissertations and Theses database. (UMI No. 3241934)

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

Naftulin, D. H., Ware, J. E., Jr., & Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Education*, 48, 630-635.

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, 249-261. doi:10.1177/0022429413497234

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.

Perry, R. P., Abrami, P. C., & Leventhal, L. (1979). Educational seduction: The effect of instructor expressiveness and lecture content on student ratings and achievement. *Journal of Educational Psychology*, 71, 107-116.

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. (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

Ware, J. E., Jr., & Williams, R. G. (1975). The Dr. Fox effect: A study of lecturer effectiveness and ratings of instruction. *Journal of Medical Education*, 50, 149-156.

Ware, J. E., Jr., & Williams, R. G. (1977). Discriminant analysis of student ratings as a means for identifying lecturers who differ in enthusiasm or information-giving. *Educational and Psychological Measurement*, 37, 627-639.

Williams, R. G., & Ware, J. E., Jr. (1977). An extended visit with Dr. Fox: Validity of student satisfaction with instruction ratings after repeated exposures to a lecturer. *American Educational Research Journal*, 14, 449-457.

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., & Madsen, K. (1998). The evaluation of teaching in choral rehearsals. *Journal of Research in Music Education*, 46, 469-481. doi:10.2307/3345486