The Effect of Behavioral Contracting on Preservice Elementary Teachers’ Performance Achievement on the Soprano Recorder

John Okley Egger, *University of Kentucky*
D. Gregory Springer, *Boise State University*
Lori F. Gooding, *University of Kentucky*
The Effect of Behavioral Contracting on Preservice Elementary Teachers’ Performance Achievement on the Soprano Recorder

**The Recorder as a Teaching Tool**

Learning to play the soprano recorder is a desired learning outcome across many collegiate educational settings, including elementary music methods courses. The recorder is frequently used as an instructional tool in many elementary classrooms, particularly between the third and fifth grade (Campbell & Scott-Kassner, 2014; Marshall & VanHaaren, 2006; Reynolds & Gottschalk, 2009). As a result, topics related to recorder performance and pedagogy are often included in textbooks for preservice music teachers (e.g., Campbell & Scott-Kassner, 2014). Additionally, these topics are also included in textbooks for preservice elementary classroom teachers as vehicles for gaining musical skills (Anderson & Lawrence, 2010; Fallin & Tower, 2011; Herrold, 2001; Hoffer, 2005; Lindeman & Hackett, 2010; Rozmajzl & Boyer, 2006; Winslow, Dallin, & Wiest, 2001).

Reynolds and Gottschalk (2009) claim that “the recorder is far more than an instrument whose only purpose is to serve as a tool to help third and fourth graders begin to read melodic notation” (p. 34). In fact, playing the recorder offers both musical and nonmusical benefits. Recorder playing may provide reinforcement of (a) music-reading skills (Herrold, 2001; Winslow, Dallin, & Wiest, 2001), (b) improvisational skills (Rozmajzl & Boyer, 2006), (c) composition skills (Stephenson, 2012), (d) recognition of the elements of music (Campbell & Scott-Kassner, 2014; Marshall & VanHaaren, 2006), and (e) finger dexterity and coordination (Campbell & Scott-Kassner, 2014). The recorder can even be adaptively modified for students with physical disabilities, making it a highly accessible musical instrument (Darrow, 2012).
Musical Skills, the Recorder, and Elementary Education Majors

Researchers have suggested that preservice elementary teachers do not feel comfortable teaching music as a subject but support its integration in other academic subjects (Hash, 2010). Preservice elementary teachers commonly lack confidence in teaching skills related to musical performance, even though certain aspects of music teaching involve singing or playing instruments. Performing on the soprano recorder in undergraduate courses could better prepare preservice elementary teachers by enhancing their confidence to teach music and by providing them opportunities to advance their musical skills. Furthermore, the experience of playing the recorder permits these preservice teachers to continue to develop their own musical skills along with their students (Winslow, Dallin, & Wiest, 2001).

Achievement related to musical skills has been reported to predict practicing elementary teachers’ uses of music activities in their classrooms (Kinder, 1987), a finding that highlights the important role of musical skill development in courses for preservice elementary teachers. Gauthier and McCrary (1999) surveyed instructors who teach music methods courses for elementary education majors and found that 70% of respondents include recorder instruction in their courses for musical skill development. Sample course syllabi (Lewis, 2002) support this finding, suggesting that the recorder is a common component of music methods courses for both future elementary teachers and future music teachers.

Behavioral Contracting as a Teaching Strategy

While playing the recorder in the elementary classroom can have both musical and nonmusical benefits, motivating students to improve recorder technique can be difficult at times because it requires concentration and practice (Howe & Sloboda, 1991). One technique that has been used in various settings to modify behavior is behavioral contracting. A behavioral contract,
also known as a contingency contract (Lee & Axelrod, 2005), is “an agreement written between two or more persons (e.g., teacher and student, therapist and client) that states behaviors the involved parties will perform and the consequences that will result” (Wolfe, 1987, p. 44). The fundamental reason for using contracts is that individuals are usually more inclined to change their behavior if the target behavior and reinforcement are mutually agreed upon between both parties. Behavioral contracts function within a theoretical framework of cognitive behavior modification, which is described as a combination of operant learning with more contemporary approaches such as social learning theory and cognitive theory (Bos & Vaughan, 2002).

Cognitive behavioral strategies have been recommended as effective tools for managing student behavior in elementary classrooms, particularly for students with behavioral disorders (Courtney, 2012; de l’Etoile, 2005).

As advised by Lee and Axelrod (2005), effective behavioral contracts should (a) be written in a positive form (i.e., specifying a reinforcement for achieving some behavioral target, rather than specifying a punishment for not achieving that target), (b) include small-scale reinforcers as a reward for reasonably attainable goals, and (c) specify terms that are clear to both the teacher/therapist and the student/client. Well-defined terms are essential because the contracts function as a firm agreement between both parties and can later be revisited when questions arise or in cases of uncertainty (Madsen & Madsen, 1998).

The efficacy of behavioral contracting has been supported by an abundance of research throughout many different educational and therapeutic domains. Behavioral contracts have been found to be effective (a) in treating alcoholism (Miller, 1972), (b) in promoting and managing physical exercise (Wysocki, Hall, Iwata, & Riordan, 1979), (c) in managing the behavior of juvenile delinquents (Stuart, 1971), (d) in promoting physical rehabilitation participation of
patients in a burn care facility (Simons et al., 1978), (e) in improving dietary compliance of chronic renal dialysis patients (Keane, Prue, & Collins, 1981), (f) in facilitating smoking cessation (Bowers, Winett, & Frederiksen, 1987), (g) on students’ grades (Williams & Anandam, 1973; Poppen & Thompson, 1971), and (h) in improving the performance of students with emotional and behavioral difficulties in the elementary school setting (Ruth, 1996).

Though the use of behavioral contracting has long been reported to be effective as a mechanism of behavior change, few studies have examined the effects of contracting in musical settings. Wolfe (1987) found that the use of behavioral contracts resulted in increased weekly practice time among piano students when compared to a baseline condition. In another study, Gooding (2009) found that the use of behavioral contracting had a significant impact on undergraduate students’ guitar performance skills. In Gooding’s study, students in a behavioral contracting condition generally showed decreases in error rates and increases in tempo compared with baseline measures.

Given that musical skills achievement is predictive of elementary teachers’ use of music activities in the classroom (Kinder, 1987), the purpose of the current study was to examine the effect of behavioral contracting on preservice elementary teachers’ performance achievement on the soprano recorder. Preservice elementary teachers’ performance achievement on the recorder was chosen because it is a common musical skills component of many music methods courses. In this study, achievement was measured as the tempo and number of errors performed on three music excerpts written for the soprano recorder. Questions posed were the following: (1) What is the effect of behavioral contracting on the performed tempo of an assigned performance selection? (2) What is the effect of behavioral contracting on the number of errors performed on an assigned performance selection?
Method

Participants

Participants \((N = 60)\) were elementary \((n = 56)\) and special education \((n = 4)\) majors from a large Southern University. They were sampled from four class sections of a music methods course for preservice elementary classroom teachers, which were taught by two of the investigators. The course was a required component of the elementary education and special education degree plan at the university and a prerequisite course for admission into the Teacher Education Program. Graded components of the course included weekly assignments, a midterm exam, microteaching experiences, a concert review, and a final portfolio. Performance on the soprano recorder was also included as a musical skills component of the course. While there were 62 students enrolled in the four class sections, only those students who attended class on the dates of the three testing sessions were included in the sample, resulting in a final sample size of 60. In terms of degree classification, the sample included freshmen (8.1%), sophomores (53.2%), juniors (30.6%), and seniors (8.1%). The majority (91.9%) of participants were female, and the mean age was 20.15 years \((SD = 1.69)\).

Participants completed a brief questionnaire that collected demographic data. These data suggested that participants’ prior musical experiences were limited. Although over half (69.4%) of the participants reported previous experience with recorder playing in elementary or middle school, the number of years of playing experience \((M = 1.20, SD = 1.08)\) was minimal. In addition to recorder playing experience, participants reported limited years of experience performing in school band \((M = 1.58, SD = 2.61)\), orchestra \((M = .18, SD = .61)\), and choir \((M = 1.43, SD = 2.12)\).

Research Design
The design of this study was based on an earlier investigation of the effects of behavioral contracting in an undergraduate music setting (Gooding, 2009). While Gooding’s study used a multiple baseline design, the present study utilized an ABA withdrawal design (Gay, Mills, & Airasian, 2009) due to a unique class scheduling arrangement. The four class sections met three times per week (50 minutes on Mondays and Wednesdays; 110 minutes on Fridays). On the Friday meetings, however, the first two classes and the last two classes met concurrently in a laboratory-style large classroom setting. Therefore, it would not have been possible to independently implement a behavioral contract with single class sections due to their joint meetings at the end of each week. As a result of this scheduling system, a withdrawal design was chosen in which all participants received the same baseline-treatment-baseline schedule.

The study was divided into three different phases, each consisting of two-week intervals, for a total of six weeks. At the conclusion of each of the three phases (T1, T2, and T3), data were collected through video-recordings of individual participants performing an assigned excerpt. Each participant’s tempo and number of errors performed were scored as data points for all testing sessions. For this study, the first phase (T1) served as a baseline measure. During the second phase (T2), contracts were implemented with all participants. Finally, during the final phase (T3), the contract was withdrawn in a final baseline measure. At the conclusion of the study, participants provided a written response to the following open-ended question: “How did the behavioral contract affect your ability to play the recorder?”

**Procedure**

Institutional Review Board approval was granted prior to the beginning of the study, and all participants provided informed consent by signing an approved consent document. To evaluate students’ performance achievement on the soprano recorder, participants played in three
“recorder check-ups,” which occurred at the conclusion of each two-week phase. At each recorder check-up, participants played an assigned excerpt, which was composed of the first seven notes learned in the class—B4, A4, G4, C5, D5, E4, D4. To eliminate practice effects associated with performing the same excerpt at each recorder check-up, three different excerpts were created using a randomly ordered set of notes. To create these excerpts, the investigators used the first seven notes listed above, and each note was included twice. Each excerpt was notated as quarter notes, which resulted in excerpts of equivalent length (14 quarter notes written in common time). The investigators randomly assigned the order of notes for each excerpt such that no repeated adjacent pitches were observed. Because the excerpts used the same rhythm and same set of notes arranged in a random order, the three excerpts demonstrated no systematic differences in difficulty. The three excerpts used for each recorder check-up are shown in Figure 1.

Excerpts were distributed to each participant on the first day of each phase. All participants were given two weeks to practice each excerpt before performing the recorder-check-up. The instructors for the course provided instruction on the excerpts during class sessions, and all classes were given the same amount of instruction time on each excerpt (5 minutes of instruction for each Monday and Wednesday class; 10 minutes of instruction for each Friday class). A metronome was used at every class session and recorder check-up to establish a steady tempo. During class time, a wide range of tempos was used to challenge participants of varied skill levels. Instruction time in all classes included a mix of teacher-led instruction on the excerpts to be played for the recorder check-ups and small group, student-led instruction on chosen excerpts from the class recorder book (Newman, 2006). Both instructors followed parallel lesson plans for recorder instruction. Recorder performance constituted 10% of the
students’ final grade in all class sections, and students were given completion grades for completing the three recorder check-ups. Completion grades were used to accommodate students who had limited prior musical performance experience and to reduce confounding threats to the effect of behavioral contract.

At each recorder check-up, participants performed individually for the course instructor in a randomly determined order. Before participants played the assigned recorder checkup, they announced their chosen tempo to the instructor. The instructor then set the metronome at the announced tempo, and participants performed the excerpt with the audible metronome. All sessions were video recorded for future scoring.

The first (T1) and third (T3) phases served as baseline measures for the study. During the second phase (T2), however, a behavioral contract was implemented (see Appendix A), which was based on the contract developed by Gooding (2009). As recommended by Lee and Axelrod (2005), the contract was constructed to provide reinforcement for successfully achieving a short-term goal. For this study, the goal was set forth in terms of the tempo and the number of errors performed. In each class, both the reward and the target goal were chosen by class consensus. For the reward, participants could choose from the following: one free absence, drop lowest grade, exemption from one concert review, or individual recorder instruction. All class sections chose the concert review exemption as their reward.

In order to outline the requirements of the behavioral contract, participants in each group were also asked to choose a percentage of tempo improvement by class consensus and instructor approval. The chosen percentages of improvement for each group were the following: 40% for Group 1, 30% for Group 2, 25% for Group 3, and 30% for Group 4. As stated in the contract terms, participants were expected to perform the excerpt for the second recorder check-up at or
above the goal tempo with fewer than four errors to earn the reward. Errors were operationally defined as follows: (a) wrong notes, (b) correct notes performed out of phase with the metronome, (c) restarts, and (d) overblown harmonics (“squeaks”). While the reward and target goal were chosen by group consensus, the rewards were given to participants individually, based on their own performance.

Because two of the investigators served as instructors for the course, a trained observer blind to condition scored the number of errors while viewing the video-recorded performances. This additional control was implemented to avoid investigator bias. The observer was a graduate student in music education who had elementary school teaching experience and training in Orff Schulwerk.

To evaluate the reliability of scoring, a reliability observer also scored the number of errors performed on a random selection of 20% of the video recordings. The random selection included an equal number of videos from each of the recorder check-ups. The reliability observer was a graduate music therapy student who had completed training in Orff Schulwerk and applied behavior analysis. Interobserver reliability, calculated as a Pearson product-moment correlation coefficient, was found to be .92. For all subsequent data analyses, an alpha level of .05 was chosen as the criterion for statistical significance.

**Results**

**Tempo**

At each testing session (T1, T2, and T3), participants chose their own tempo and performed the excerpts with an audible metronome pulse. In terms of performed tempo, means (notated as beats per minute) increased at each testing session. Descriptive statistics for performed tempo across the three testing sessions are provided in Table 1. A repeated measures
analysis of variance (ANOVA) was conducted to examine differences in performed tempo among the three testing sessions. Because Mauchly’s test indicated that these data did not meet the assumption of sphericity, $W = .58, \chi^2(2) = 31.45, p < .001$, a Greenhouse-Geisser correction was applied, which resulted in an adjustment to the degrees of freedom for the within-subjects effect. Results of the omnibus ANOVA test indicated a significant difference in performed tempo among the three testing sessions, $F(1.41, 83.19) = 451.47, p < .001$, partial $\eta^2 = .88$, which represents a very large effect size (Cohen, 1988). Follow-up pairwise comparisons (with a Bonferroni correction to control for inflated Type I error) showed significant differences among all three testing sessions ($p < .001$), indicating that the tempo increases between T1 and T2, between T1 and T3, and between T2 and T3 were all statistically significant. Figure 2 illustrates the significant increases in participants’ performed tempo across the three testing sessions.

**Errors**

Aside from the tempo performed at each testing session, the number of errors performed was also scored as a dependent variable. As mentioned above, errors were operationally defined as follows for the purposes of this study: (a) wrong notes, (b) correct notes performed out of phase with the metronome, (c) restarts, and (d) overblown harmonics (“squeaks”). Each of the error types described above was counted as one error. Thus, all participants started with an error score of zero, and one point was added for each error performed. Each type of error could theoretically occur multiple times during a testing session (e.g., multiple restarts), and errors were scored each time they occurred. As a result, a relatively large range of performed errors was observed. Throughout the duration of the study, observed errors at a single testing session ranged from 0 to 28.
As shown in Figure 3, the mean number of errors decreased from T1 to T2, but it increased from T2 to T3. Descriptive statistics for number of errors performed are provided in Table 1. A repeated measures ANOVA was conducted to examine differences in the number of errors performed across the three testing sessions. Again, a statistically significant Mauchly’s test indicated that these data did not meet the assumption of sphericity, $W = .82, \chi^2(2) = 11.69, p = .003$, so a Greenhouse-Geisser correction was applied. Results of the repeated measures ANOVA indicated a significant difference in the number of performed errors, $F(1.69, 99.79) = 22.03, p < .001$, partial $\eta^2 = .27$. This also represents a large effect size (Cohen, 1988). Follow-up pairwise comparisons (with a Bonferroni correction) showed significant mean differences between T1 and T2 ($p < .001$) and between T1 and T3 ($p < .001$). No significant differences were observed between T2 and T3.

Discussion

The current study investigated the effect of behavioral contracting on preservice elementary teachers’ performance achievement on the soprano recorder. As noted in the literature review, the use of the recorder in the classroom provides numerous benefits for elementary students. Therefore, recorder performance and pedagogy are commonly included in textbooks for preservice elementary classroom teachers (cf. Anderson & Lawrence, 2010; Fallin & Tower, 2011; Hoffer, 2005; Lindeman & Hackett, 2010). Because some students are less motivated to succeed at certain tasks requiring goal-oriented concentration while practicing (Howe & Sloboda, 1991), results from this study suggest that behavioral contracts may be one option to provide an additional motivational tool for preservice classroom teachers based on achievable short-term goals and incentives.

Limitations
Results from this study should be interpreted with some limitations in mind. First, due to the scheduling of the course, the investigators used a withdrawal design (ABA) to accommodate the simultaneous class meetings at the end of each week. While it would be beneficial to bring back an additional contracting phase at the end of the study (ABAB), this was not feasible due to the placement of a school holiday, which would interrupt the final contracting phase by one week. Thus, in an effort to control for threats to internal validity, the withdrawal design was chosen as an alternative to a multiple baseline design or other extended behavioral design.

Second, these findings were collected from a sample of preservice teachers from one university, so the applicability of these findings to other universities is unknown. Finally, because the class meetings were led by two of the investigators, the results of the study could have been influenced by teacher effects, although efforts were made to diminish these effects. To reduce these threats, classroom recorder instruction was standardized such that instructors used the same amount of class time working on the recorder excerpt at each class session, the same testing dates, parallel lesson plans, and a consistent use of the metronome at all class sessions for each of the classes.

**Summary of Results**

Tempo and number of errors performed were the dimensions of performance achievement that were measured in this study. With each phase, participants demonstrated a statistically significant increase in performed tempo, however the largest increase was observed during the contracting phase (T2). Even after the contracting phase, performed tempos continued to increase after the contract was removed, and the difference was statistically significant. These results are consistent with the basic premise of a contract, i.e., improvement is maintained after the contracting period ends (Miller, 1972). Other factors could have influenced these results, such as the natural improvement of performance skills due to increased practice time and
experience playing the recorder, but the decision to use three different excerpts was made to reduce the influence of these factors.

In terms of the number of errors performed, a different trend was observed. From the first baseline phase (T1) to the contracting phase (T2), a statistically significant decrease in the number of performed errors was observed. A small non-significant increase in errors was observed between the contracting phase (T2) and the final baseline phase (T3), but these observed differences could simply be due to chance.

Participants’ responses to the open-ended question (“How did the behavioral contract affect your ability to play the recorder?”) provided tentative explanations about how they perceive the effectiveness of the behavioral contract. Proponents of the behavioral contract claimed that the contract positively affected their performance skills on the recorder (a) by providing greater motivation to practice, (b) by offering a specific short-term goal and incentive, and (c) by encouraging them to practice for higher skill or achievement levels. For example, one participant responded, “I practiced a lot more for [the second] check-up than I did for the others. I made sure that I could play confidently and without errors at a much faster tempo than I did previously.” Another participant explained that “the behavioral contract affected my ability to play the recorder by giving me motivation to always go above and beyond what is required. Although I could have gone with the minimum, I went further because I felt I needed to earn the reward.” When referring to the effect of the behavioral contract, one participant claimed that “… my motivation to do well was definitely stronger” and candidly admitted, “I actually practiced for [the second] one! 😊 The goal was clear, and I felt that it was reasonably attainable, plus I really wanted the reward!”
Participants who opposed the behavior contract claimed that the contract made them feel more nervous and feel greater pressure in an effort to play perfectly. For example, one participant responded, “I think the behavioral contract made me a little more nervous. I felt like I was under more pressure, whereas the other recorder check-ups were more relaxed and I felt more comfortable.” Another participant attributed a negative effect of the contract “…because the stakes were higher and I was nervous about it. I didn’t do as well but I also think that it had a lot to do with the camera.”

Implications and Suggestions for Future Research

After the data were collected, the instructors discussed the practical implications of behavioral contracting with students in each of the classes. Discussions included how to effectively implement behavioral contracts within elementary classroom settings, especially when teaching students with special learning needs (Courtney, 2012; de l’Etoile, 2005; Ruth, 1996). These discussions primarily focused on how behavioral contracting can be used as a teaching tool outside of musical performance contexts (e.g., homework grades, class participation, tests, and attendance), which were intended to encourage students to make a practical transfer to their future classrooms.

Overall, the single implementation of a behavioral contract appeared to positively impact participants’ performance achievement on the soprano recorder. Participants’ tempo and number of errors performed improved as a result of the behavioral contract. These findings are consistent with prior research, which reports positive effects of behavioral contracting on musicians’ practice time (Wolfe, 1987) and performance skills (Gooding, 2009).

Because the results of this study support the use of behavioral contracting among preservice elementary teachers, future studies should examine the effects of behavioral
contracting on preservice music teachers or other groups of musicians (e.g., college musicians in applied lessons, music majors learning secondary instruments, secondary school instrumentalists, etc.). Other related variables could also be studied in future investigations in conjunction with performance achievement, including amount of weekly practice time, distribution of practice time within practice sessions, and content of practice sessions. Future investigations that study the effects of behavioral contracting on both performance achievement and practice behaviors would be particularly beneficial. As suggested by the results of the present study and prior research, the use of behavioral contracting can create a positive impact on student’s performance achievement.
References


Table 1

Descriptive Statistics of Tempo and Error Scores across Testing Sessions

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (T1)</td>
<td>53.16</td>
<td>6.36</td>
<td>40</td>
<td>76</td>
</tr>
<tr>
<td>Contract (T2)</td>
<td>72.74</td>
<td>9.19</td>
<td>56</td>
<td>100</td>
</tr>
<tr>
<td>Baseline (T3)</td>
<td>79.43</td>
<td>11.05</td>
<td>63</td>
<td>120</td>
</tr>
<tr>
<td>Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (T1)</td>
<td>6.61</td>
<td>5.99</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>Contract (T2)</td>
<td>2.69</td>
<td>3.90</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Baseline (T3)</td>
<td>3.85</td>
<td>5.31</td>
<td>0</td>
<td>28</td>
</tr>
</tbody>
</table>

Note: Tempo means are expressed as beats per minute. Error means are expressed as a frequency count.

Figure 1. Excerpts Performed at each Recorder Check-Up
Figure 2. Performed Tempo across Three Testing Sessions

Figure 3. Number of Errors Performed across Three Testing Sessions
Appendix A

Behavioral Contract

I understand if I achieve a _____% or higher improvement in playing the second recorder check-up with fewer than four (4) errors, then I will receive the following reward: _______________________________________________. *Reward Options:* 1 free absence, drop lowest grade, exemption from concert review, or individual recorder instruction.

**Contract Stipulations:**
Percentage improvement to be chosen by students from an instructor suggested range by simple majority. Improvement is individual and is based on percentage. For example, if you start out at 30 beats per minute and the required improvement is 50%, then in order to receive the reward you would need to perform the excerpt at 45 beats per minute at the end of the contract period. If you start at 60 beats per minute with the same percentage, then you would need to perform the excerpt at 90 beats per minute at the end of the contract period.

Improvements will be determined in class during videotaped recorder check-ups. Tempos for the recorder checkup will be chosen by the students. The student will inform the instructor what tempo he/she would like to go and then play the melody with a metronome while being videotaped. Errors will be determined based on four criteria: (1) wrong notes, (2) wrong time (out of phase with metronome), (3) restarts, and (4) squeaks.

***N.B. Remember that your grade for the recorder check-up is based on completion. The reward option, however, is based on your improvement in tempo and the number of errors you make.***

Reward to be chosen from instructor/peer suggestions by students through simple majority.

Video 1: Tempo: _________________
Video 2: Tempo: _________________
Video 3: Tempo: _________________ % Improvement of _________________

Name (signed): ________________________________________________________________

Name (printed): _______________________________________________________________