THE EFFECT OF GENERAL-CASE TRAINING, INSTRUCTIONS, FEEDBACK, AND REHEARSAL ON THE ACQUISITION OF MUSIC SIGHT-READING BY ADVANCED FLUTE STUDENTS

by

NANCY ELLEN DIB

A dissertation submitted to the Graduate Faculty in Psychology
in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York
2010

This manuscript has been read and accepted for the
Graduate Faculty in Learning Processes Psychology in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date

Dr. Peter Sturmey
Chair of Examining Committee

Date

Dr. Maureen O'Connor
Executive Officer

Dr. Bruce Brown
Dr. Emily Jones
Dr. Bertram Ploog
Dr. Parsla Vintere

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECT OF GENERAL-CASE TRAINING, INSTRUCTIONS, FEEDBACK, AND REHEARSAL ON THE ACQUISITION OF MUSIC SIGHT-READING BY ADVANCED FLUTE STUDENTS

by

Nancy Ellen Dib

Advisor: Professor Peter Sturmey

Sight-reading music enables the performance of music that has not been previously learned. Without sight-reading skills, required behavior (e.g., learning new work, performing new music, and passing musical exams) is equivalent to learning a piece of standard repertoire. Therefore, all students should learn to sight-read. To date, no research has been done on the use of applied behavior analysis for teaching students how to improve music sight-reading. Sight-reading may be more efficiently taught if it is approached by planning for generalization of music-related behavior in music education. Therefore, the current study taught advanced flute students to improve their sight-reading skills with a treatment package that included general-case training, instructions, feedback and rehearsal. This study used a multiple-baseline-across-subjects research design for three advanced flute students during their regular lessons. There was a systematic decrease in sight-reading errors as treatment was introduced across subjects. Note errors and rhythm errors decreased by an average of 10% and 42% respectively. Frequency of repetitions and hesitations decreased by an average of 7 and 2 respectively. Therefore, the training package was effective in improving music sight-reading. Future research should investigate the use of general-case training and/or behavioral skills training in other flute-playing behavior, as well as in the teaching of other instruments. Future research should also investigate

the components of the current package individually to determine if they would be as effective separately.

Acknowledgments

This dissertation would not have been possible without the help and support of many very important people in my life. First, I would like to thank my advisor, Dr. Peter Sturmey. This dissertation would not have been possible without your guidance, knowledge, and feedback. You have been a wonderful teacher and have encouraged me to become the researcher, behavior analyst, and public speaker that I am today. Thank you for meeting with me on weekends, providing feedback and suggestions so quickly, and most of all for being as excited about my research as I am.

To my parents, Gil and Nancy Breines, it is not possible for me to come up with the words to express how thankful I am to have you as my parents. Thank you for your unconditional support and encouragement. You have been wonderful role models to me all of my life. You taught me to have a strong work ethic and to never stop reaching for my goals. Thank you for everything you have done and continue to do for me.

And finally, thank you to my husband, Matthew Dib, for all of your love and support throughout my graduate career. I could not have finished this dissertation without your technical support and assistance with scoring hours of data. Thank you for your constant encouragement and pushing to "get things done." You kept me working even when I felt like stopping. I could not have asked for a better husband.

Table of Contents

List of Figures

List of Tables

Introduction

Sight-reading is an important skill for all musicians to have. This is because sight-reading enables musicians to play any new music without prior practice, such as duets and ensembles, accompaniment for other musicians, "requests" for family members, friends, and patrons in commercial settings. Furthermore, sight-reading is required for state exams and auditions. Music students report that sight-reading is the most difficult test to pass because it is the only part of performing that they can not prepare for directly. Typical training often fails to prepare musicians for tests of sight-reading. It is important for musicians to sight-read music because the skill greatly improves musical performance. The ability to sight-read enables students to play new music without getting stuck on reading music notes. This allows students to concentrate on putting in musical expression, tone, balance, etc. It is for these reasons that behavior-analytic research should address sight-reading.

Much research has demonstrated the usefulness of applied behavior analysis in teaching a variety of skills to subjects of all ages. For example, Dib (2007) reviewed research that attempted to use a behavioral approach to teaching acquisition, modification, and maintenance of music-related behavior. This systematic review of the literature found only 24 behavioral studies that addressed music related behavior and only five met all of Poulson and Nunes' (1988) six criteria for experimental analytic studies. Of the five studies that met all criteria (Dib & Sturmey, 2007; Eisenstein, 1976; Salzberg & Greenwald, 1977; Salzberg & Salzberg, 1981; Wolpow, 1976), two investigated the effect of contingency manipulations (Salzberg & Greenwald, 1977; Wolpow, 1976) and three investigated the effects of behavioral skills training packages (Dib & Sturmey, 2007; Eisenstein, 1976; Salzberg & Salzberg, 1981).

Salzberg and Greenwald (1977) used a token system to improve attentiveness and punctuality in two string instrument classes. Twenty-one seventh-grade students enrolled in the second year of string class in violin, viola, or ‘cello participated. The dependent variables were student attentiveness and punctuality and the independent variable was the implementation of a token system. The researchers used a multiple-baseline design to determine the effectiveness of the token system. During baseline, 46% and 50% of students were punctual in the two classes, whereas during intervention, 100% of students in both classes were punctual. The mean percentage of on-task behavior increased from 74% of intervals during baseline to 92% of intervals when the token system was in effect for one class and from 54% of intervals during baseline to 85% of intervals during treatment for the second class. Therefore, the token system changed student punctuality and on-task behavior during music instruction. The authors concluded that antecedent reminders were ineffective, but contingent points were effective in controlling attendance and on task behavior during music lessons. Although the results demonstrated the effectiveness of the token system, the dependent variables of the study were not specific to musicians. That is, attentiveness and punctuality are two skills that can be applied to any task.

Wolpow (1976) also found that contingency manipulations affected music-related behavior. Wolpow investigated the effect of contingent approval and disapproval on musical on-task and performance behavior. Seven adults with profound mental retardation in a music therapy rhythm band at a residential institution participated. The dependent variables were musical on-task and performance behavior and the independent variables were contingent teacher approval and disapproval. The author defined experimenter interactions with students as *academic approval*, contingent approval of an academic performance behavior; *social approval*, contingent approval of on-task behavior; *academic disapproval*, contingent disapproval of a performance behavior; and *social disapproval*, contingent disapproval of off-task behavior. A multiple-baseline-across-responses-research-design showed that the mean proportion of students' on-task increased from 54% to 78% and the mean performance scores increased from 28% to 50%. High rates of approval for appropriate on-task and performance behavior resulted in a significant increase in both types of behavior. Of the two dependent variables, only one was specific to musicians since on-task behavior can be measured for most skills. Musical performance behavior, however, is a response specific to musicians because it requires playing an instrument.

The two foregoing experiments suggest that contingencies are effective in modifying on-task, off-task, and attendance behavior in music lessons and that contingencies are more effective than antecedent stimuli alone. Since neither of these two experiments manipulated contingency independent of the presentation of consequences, future studies should include both non-contingent and contingent experimental conditions. The lack of reported effects of antecedents may merely reflect the lack of variety of antecedents investigated. Therefore, future studies should investigate both a variety of other antecedents and how to establish stimulus control of music behavior.

Three studies used packages of instructions, modeling and other antecedents, rehearsal, and various consequences to modify several different music-related behaviors. For example, Dib and Sturmey (2007) used verbal instruction, modeling, rehearsal, and feedback, to teach three beginning flute students aged 8-9 years correct posture during flute lessons. The dependent variable was a measure of student posture and the independent variable was the behavioral skills training package. An experimenter videotaped the students during their regular 30-minute flute lesson once a week for three to four months. The authors used a multiple-baseline-across-subjects experimental design to demonstrate experimental control. Independent observers collected data using 10-second partial interval time sampling. During baseline, observers collected data on student posture and the instructor modeled correct posture. During treatment, the experimenter implemented the behavioral skills training package. Observers also took generalization data on probes when students practiced musical scales. Probes, without treatment in place, took place throughout the study to test for generalization. All three students' posture improved from 0% during baseline to nearly 100% after training for all sessions, during generalization probes, and after a one- to two-month follow-up. Therefore, the training package was effective in the acquisition, generalization, and maintenance of correct posture for flute playing. The authors also noted that, although the teacher always played with correct posture and all of the students were able to describe the teacher's posture, none of the students played with correct posture during baseline, suggesting that modeling alone was ineffective to teach proper posture to beginning music students.

Eisenstein (1976) used behavioral skills training to teach music symbol names. He described his intervention as "successive approximations." The intervention is here considered behavioral skills training, as it included manual prompting, fading, and praise. The dependent variable was naming music symbols and the interventions were prompting plus reinforcement or reinforcement alone. Eisenstein used a combined reversal and multiple-baseline design to examine the effect of clues, a flashing light, and no clues for naming music symbols, which he presented on cards. Each subject participated in two consecutive, 7-minute baseline or treatment sessions per day on five consecutive days. During baseline, Eisenstein presented subjects with a music symbol picture card and asked them to match the appropriate verbal name card. During treatment phase 1, Eisenstein modeled correct responses and provided discriminative stimuli by moving the subject's hand toward the correct card, immediately approving and flashing a feedback light if the subject hesitated near the correct card. In Phase 2, Eisenstein only reinforced correct responses with verbal approval and a flashing light, omitting modeling and cues. During baseline, participants correctly named rests during 10.67% of trials, correctly named note values during 29.33% of trials, and correctly labeled dynamic markings during 32.00% of trials. During treatment, the corresponding data were 89.00%, 82.67%, and 89.33%, respectively. Therefore, the treatment package was effective for teaching music symbol names to all subjects.

Finally, Salzberg and Salzberg (1981) used corrective feedback, prompting, and praise, to teach correct left-hand position to five beginning string students in fourth-, fifth-, and sixth-grade, aged from 9-12 years. Four subjects played violin and one played viola during regularly scheduled string classes that met for 45 minutes twice a week. The dependent variable was the number of seconds the student's left-hand position was incorrect during two, one-minute playing periods in each class. The independent variable was the behavioral skills training package. Researchers collected probe data to measure generalization of the behavior to another period of class when the target behavior was not specifically taught and collected generalization probe data for one minute of playing time during the music reading segment of class. The researchers used a multiple-baseline-across-subjects experimental design. The phases were Baseline, Interventions 1, 2, and 3. Intervention 1 was corrective feedback, Intervention 2 was prompting and praise with no corrective feedback, and Intervention 3 was prompting and praise for an increased period of time and with an increased number of praise statements with no corrective feedback. Researchers assigned the first three subjects to baseline and all interventions; the fourth subject to baseline and interventions 1 and 3; and the fifth subject to baseline and intervention 3 only. During baseline all subjects emitted incorrect left-hand position during 100% of intervals. There was a decrease in incorrect left-hand position for Interventions 1, 2, and 3 to an average of 71.0%, 72.3%, and 42.0% of intervals with incorrect hand position during training and 96.5%, 96.5%, and 79.2% of intervals with incorrect hand position during generalization probes, respectively. Thus, Intervention 3 resulted in the greatest reduction in intervals with incorrect responses during both treatment and generalization probes. This was the only study that compared components of the behavioral skills training package.

These five behavior analytic studies found: (a) evidence for the effectiveness of contingency-based interventions; (b) no evidence for the effectiveness of antecedent interventions, such as instructions and modeling alone; (c) three studies found evidence that behavioral skills training packages were effective; and (d) one study found that more praise was more effective than less praise. The fact than only five studies were found that used behavior analysis to address music behavior demonstrates the lack of behavior analytic research on this topic. There are many areas of music behavior for behavior analytic research to address, such as the investigation of a wider range of antecedent stimuli and ways to bring music behavior under appropriate antecedent control. Only Dib and Sturmey (2007) addressed stimulus generalization and only Salzberg and Salzberg (1981) addressed response generalization. Stimulus and response generalization are important topics because sight-reading and improvisation, which are both examples of generalization, are important aspects of music behavior. Sight-reading music requires the musician to respond to novel stimuli and both sight-reading and improvisation require the musician to make novel responses by playing notes and rhythms in novel combinations. Dib (2007) did not identify any current behavior analytic research on sight-reading. Sight-reading allows musicians to perform new music without extensive practice time. Sight-reading is a type of generalization in which the musician generalizes music behavior, such as reading music and playing notes and rhythms, from music that has already been learned to music that has never been played before. Therefore, research can approach the task of teaching sight-reading as an issue of generalization.

Generalization is the spread of the effect of reinforcement in the presence of one stimulus to other stimuli not correlated with reinforcement (Catania, 1992). Stokes and Baer (1977) conducted a review of generalization literature to describe the technology of generalization and found that training multiple exemplars was one of the most extensive areas of generalization-programming. They concluded that “sufficient exemplars” of stimuli or responses are required to program for generalization to untrained stimuli and responses. It is not clear exactly how many exemplars are “sufficient.” Therefore, many researchers have attempted to identify the number and type of exemplars necessary to promote generalization.

One such study was conducted by Hupp and Mervis (1981). The researchers taught six students with severe disabilities, ranging in age from 8 to 18 years, to label categories of objects using sign language. They taught labeling with three different types of exemplars, good, moderate, and poor, which were pictures of objects to be labeled. The researchers defined good exemplars as those that met guidelines for the best example theory, including family resemblance (better group members share large numbers of attributes), contrast set (better group members share fewer attributes with non-group members), and saliency (the degree to which attributes are recognizable). They found that training with only good (representative) exemplars resulted in generalization above chance levels, while training with both good and poor (non-representative) exemplars did not. Furthermore, the use of three good exemplars resulted in significantly more accurate generalization than the use of only one good exemplar. This research demonstrated that it is not only the presentation of multiple exemplars that leads to generalization, but that how representative the exemplars are is also important in the programming of generalization. The researchers showed that fewer exemplars may be more effective than many if only “good” exemplars are used.

Horner and McDonald (1982) compared single- and general-case training for the generalization of a vocational skill, crimping/cutting of biaxle electronic capacitors. Single-case training included training with one capacitor. General-case training included training on three new capacitors. Capacitors used for general-case training were selected by identifying the relevant stimulus characteristics of capacitors. These dimensions included head size in centimeters, lead distance in centimeters, and head shape. The authors selected 20 capacitors that sampled the full range of variation available in commercially made biaxle capacitors. The researchers tested for generalization of the vocational skill with 20 non-trained capacitors. The single-case training did not lead to improved performance on the test of generalization. Furthermore, single-case training was followed by a dramatic increase in errors on tests of generalization compared to baseline. On the other hand, after general-case training, subjects increased the number of probe capacitors performed correctly and decreased errors. The authors concluded that it is not enough to simply train on multiple exemplars, since poor training examples may contribute to learning of errors. Researchers should select the smallest number of examples needed to teach a skill that they would like to see generalized. These findings are consistent with those of Hupp and Mervis (1981) in that they demonstrate it is the quality and not the quantity of exemplars that has the greatest impact on generalization.

Similarly, Sprague and Horner (1984) compared single instance, multiple instance, and general-case training on generalized vending machine use for six students with mental retardation aged 15 to 19 years. For single instance, training occurred with a single vending machine chosen because it was close to the classrooms. Three similar vending machines, chosen because they did not sample all of possible variation of the probe machines, were used as training instruments for the multiple instance condition. The general-case training condition including training on three machines that sampled the range of stimulus and response variation in a defined class of vending machines that later served and generalization probes. The authors explained that these machines were selected using Horner and McDonald's (1982) "general case analysis" guidelines. Baseline data indicated that none of the students were able to operate vending machines prior to training. Of the three training procedures, only general-case training was effective in the acquisition of generalized vending machine use. After single instance training, students displayed small improvement in probe performance, with 5 out of 6 students correctly using only 1 of 10 probe machines. Across the six probes delivered under the multiple instance phase, only 9 of the possible 60 trials were performed correctly. The authors stated that after general-case training, all students showed substantial improvement across the probe machines, although it is not clear how many of the trials the students performed correctly. Furthermore, this improvement was immediate for 5 of 6 students.

Neef et al. (1990) also found general-case training more effective in promoting generalization than single-case training. These researchers analyzed the role of range of variation in training exemplars as a contextual variable influencing the effects of in vivo versus simulation training in producing generalized responding for using washing and drying machines. Subjects were four adults with mental retardation. Researchers selected stimuli based on an analysis of the washing machines and dryers that the subjects were most likely to encounter in the community (i.e., at the community residential facilities, the day habilitation center, and the local laundromat). Some dimensions of washer and dryer machines that the researchers identified included how to open the lid, clean the lint trap, set temperature, set cycle, set time, etc. Of the seven different washers and dryers surveyed, the researchers selected four or five of each for training and the remaining machines served as generalization probes. Researchers designed a simulated washing machine and dryer from photographs of each actual counterpart. After baseline, all subjects received single case training with one washing machine and one dryer. This was followed by general case training in which researchers trained on the pre-selected machines on a rotating basis across sessions until criterion was met on each. Researchers counterbalanced in vivo and simulation training across the two tasks, washing machine and dryer use, for the two client pairs. The range of training exemplars affected generalization errors and the use of simulated versus natural training stimuli did not. More errors occurred on untrained machines after single-case training than after general-case training.

Ducharme and Feldman (1992) conducted two studies to compare: (1) written instructions, (2) single-case training, (3) common stimuli training, and (4) general-case training for the promotion of generalization of staff skills in teaching self-care routines to clients with developmental disabilities. In Study 1, subjects were nine direct-care staff members working in a group home for adolescents and adults with developmental disabilities. The researchers divided the staff members into two groups and used a multiple-baseline-across-groups research design. After baseline, staff members received training sequentially through the four conditions (listed previously). The staff members did not reach generalization criteria until the researchers provided general-case training. Therefore, the researchers conducted Study 2 to control for possible sequence effects. In this study, seven direct-care staff members were trained using only general-case training after baseline and all staff members reached generalization criteria. These two studies demonstrated the effectiveness of general-case training in promoting generalization across clients, settings, and client programs over other commonly used staff-training procedures.

Many other studies have demonstrated the effectiveness of multiple-exemplar training in the promotion of generalization across a variety of behaviors, including teaching generalization of conversational interactions of high school students with moderate mental retardation (Hughes, et al., 1995), inducing the formation of linked perceptual classes for typically developing college students (Fields, et al., 2002), facilitating the development of arbitrary comparative relations for typically developing girls aged 4 to 5 years old (Berens & Hayes, 2007), establishing a generalized repertoire of helping behavior in children with autism (Reeve, et al., 2007), and establishing derived equivalence in an infant (Luciano, et al., 2007). Although all of these studies demonstrated the effectiveness of multiple-exemplar training, they were not all clear on the process of selecting exemplars. For example, one of the most well described sampling of exemplars was provided by Reeve et al. (2007). The authors mentioned the selection of relevant examples of target behavior by asking parents of typically developing children about their helping behavior and observing the helping behavior of typical children in classrooms at a local elementary school. After this selection procedure, the researchers used eight experimenter-defined response categories of helping. It is unclear how the researchers identified these multiple-exemplars. This seems to be a difficulty with the replication of all of the previously mentioned studies which utilized multiple-exemplar training because, as defined by its name, this procedure requires only that multiple examples be used for training. As with Stokes and Baer's

(1977) suggestion for "training sufficient exemplars," it is difficult to determine exactly what "sufficient" is. Another strategy for promoting generalization, general-case training, is closely related to multiple-exemplar training; however, the procedures for selecting examples to be trained are made explicit. According to Ducharme and Feldman (1992), general-case training requires researchers to carefully choose multiple teaching examples to sample the span of the stimulus and response variation that defines the "instructional universe" of the skill to be trained. The instructional universe includes all stimulus situations in which the response is expected to occur (Colvin & Horner, 1983). Colvin and Horner (1983) created a model for general case instruction which includes six steps: (1) define the instructional universe; (2) define the range of relevant stimulus and response variation within the instructional universe; (3) select examples from the instructional universe for training and probe testing; (4) sequence the training examples; (5) teach with the training examples; and (6) test with non-trained probe examples. Although, general-case training and multiple-exemplar are procedurally similar, the explicit description of exemplars for general-case training may make this procedure easier to replicate and interpret. Therefore, the current study used the general-case training procedure in an attempt to teach music sight-reading to flute students.

To date, general-case training has not been used to teach music sight-reading even though the skill of sight-reading is an issue of generalization. Furthermore, behavior analytic research has yet to investigate music sight-reading. In order to identify dimensions for the current investigation, the definition of sight-reading was broken down into four dimensions. Correct sight-reading was defined as playing a new line of music with respect to four properties: (a) notes, (b) rhythms, (c) articulations, and (d) dynamics. Therefore, these four properties defined the instructional universe. The flute students who served as participants in the present study had previously mastered articulations and dynamics. Therefore, the general-case training in this study consisted of training on two dimensions; notes, with the use of various musical key signature, and rhythms, with the use of various time signatures. The researcher used this procedure to expand on current research in both the field of generalization and of music behavior. More specifically, the current study investigated the effect of general-case training in combination with a behavioral skills training package that included instructions, feedback, and rehearsal on the acquisition of music sight-reading skills by advanced flute students during private music lessons.

Method

Participants

The participants were three typically developing students in a music class between the ages of fourteen and seventeen years. The experimenter chose these students because they were advanced students, identified by their previous participation in the New York State School Music Association (NYSSMA) Festival at a level 5 or 6 (as identified by the NYSSMA guidelines) and because they needed to prepare for the upcoming NYSSMA Festival, which occurs between April and June. The experimenter chose these students because (a) their score on sight-reading was 50% lower than their overall score on the judging sheet for a previous year, (b) typical teaching methods, including the completion of at least three method books in a popular series, were ineffective in perfecting sight-reading for these students, (c) they exhibited at least five hesitations and/or repetitions on five consecutive sight-reading pieces, and (d) when asked, they stated that they wanted to participate in the study to receive extra help with sight-reading.

Setting and Stimuli

All sessions took place in the regular music classroom in a private music school. The room was approximately 2.43 m x 2.43 m containing only two chairs, a music stand, and a keyboard in the back of the room. The experimenter positioned a video camera on a tripod in the rear left corner of the room next to the keyboard. The lessons were private, so only the student and the experimenter were in the room. The experimenter (principle investigator of the study) was the students' regular flute teacher. Each student sat on a chair next to the experimenter with a music stand directly in front of him/her. The experimenter created all music using a computer program (Finale Notepad ®). She created the pieces by placing notes on the music staff to sample two octaves of notes in an unusual pattern. The experimenter presented the musical pieces only once to each student throughout the study.

Procedure

Sight-reading was defined as reading sheet music and playing it on an instrument "at sight," or at first glance, without previously practicing it. The target behaviors for this study were five types of musical errors including (a) note errors, (b) rhythmic errors, (c) repetitions, (d) hesitations, and (e) frustrated behavior. The experimenter defined a note error as playing a pitch that was different from what was written in the music score. She defined a rhythmic error as playing a note or rest longer or shorter than its written value. The researcher defined repetitions as re-playing a part or all of the measure instead of moving on to the next measure. She defined hesitations as any pause that interrupted playing through the measure and did not occur on a rest. The experimenter defined frustrated behavior as any sound or excessive body movements made while playing the pieces (i.e., saying "I can't do it" or "It's too hard" and/or sighing, grunting, or stomping a foot after making a mistake). The experimenter chose these behaviors because they are often emitted by students when they make errors while playing music. Ultimately, the experimenter did not include frustrated behavior in the study because this target behavior never occurred when the camera was taping.

Experimental Design. The experimenter used a multiple-baseline-across-subjects-experimental-design. She videotaped each student during his/her usual 30-minute flute lesson once a week for 7 months. As summer approached, the experimenter increased the duration of each student's lesson to 60-minutes for the last 2 months of the study to collect more data and help the students complete the study before the end of the school year. The experimenter and an independent observer viewed and scored the tapes. Both were familiar with reading and playing music and received a copy of the music the student played.

Data Collection. The experimenter and the independent observer collected data by marking on a data sheet the number of musical errors including note errors, rhythmic errors, hesitations, repetitions, and frustrated behavior made while playing each musical piece. They calculated the percentage of notes with errors for note and rhythmic errors by scoring on a data sheet the number of note or rhythm errors in each piece, dividing this score by the number of notes in the piece, and multiplying by 100%. They calculated repetitions by marking on a data sheet the frequency or number of times the subject replayed a section of notes in each measure. The experimenter and the independent observer scored notes that the students played incorrectly during repetitions as a note error once, but counted multiple repetitions each time they occurred. They calculated hesitations by counting the number of hesitations to determine frequency.

The experimenter had a copy of all music and marked errors on it while the student played by circling incorrect notes, marking a minus sign for notes that were too short and a plus sign for notes that were too long. In addition, the experimenter wrote the musical symbol for "pause" for hesitations, "R" for repetitions, and "F" for frustrated behavior on her copy of the music. If any error occurred more than once, she used tally marks above the letter to indicate the number of times the error occurred.

As mentioned previously, the dimensions that the experimenter used for general-case training included notes and rhythms. Dimensions for correct notes included multiple key signatures. Training included all flat (one through seven flats) and all sharp (one through seven sharps) key signatures. Dimensions for correct rhythms included multiple time signatures. Training included multiple time signatures, specifically the most commonly played time signatures (4/4, 3/4, 2/4, and 6/8) and time signatures seen less frequently by the students (2/2, 3/8, 9/8, and 12/8).

The principle investigator, who was also the students' regular flute teacher, used a matrix (see Appendix A) to create sight-reading stimuli to sample all of the possible key signature and time signature combinations. She created all 120 of the sight-reading pieces, including 15 training pieces and 105 baseline/follow-up pieces. All pieces were 10 measures long. To control for difficulty, all pieces contained the same number of each type of note (whole, half, quarter, eighth, sixteenth, and thirty-second notes) for each time signature. To further control for difficulty, the experimenter randomized the pieces given to the students to play by assigning numbers to all of the pieces and then using a computerized random number generator to determine the order of piece presentation. The experimenter determined this order for each participant prior to the start of the study and no pieces were repeated.

The experimenter required all students to record all of their practice time in a practice log and turn in the log at the beginning of their lesson each week to determine whether any improvement during the study could have been due to an increase in practice time from baseline to treatment. She instructed the students to record the date of each practice session and time at the start and end of each practice session. The experimenter told students during each weekly lesson that sight-reading is important and reminded them to practice it (i.e., "Sight-reading is an important skill, you should practice it regularly.").

Baseline. During the baseline phase, the experimenter collected data on musical errors (defined previously) made by the students while playing pieces of music. She required students to play a random selection of the researcher-created sight-reading stimuli. The experimenter had a copy of all music and marked errors on it while the student played, as described previously.

She provided general feedback within 10 seconds of completion (e.g., "Good try, let's try another one."). The students played each piece only once and none of the baseline pieces were repeated at any point during the study. The experimenter collected data on each musical piece during the music lesson once a week until the data were stable. She defined stability as a steady or upward trend in sight-reading errors for at least 3 consecutive sessions.

Treatment. During treatment, the experimenter implemented a training procedure. She required students to play the fifteen treatment pieces in a random order and trained on these pieces using a behavioral skills training package including, verbal instruction, feedback, and rehearsal. For the verbal instruction, the experimenter told students what correct sight-reading is and what to look for when preparing to play a new piece of music (e.g., "Correct sight-reading includes playing the correct notes, rhythms, articulations, and dynamics. Remember to look at the time and key signatures before playing."). The experimenter then asked the student to sight-read the piece of music. She identified errors on the music in the same way as during baseline and this rehearsal was immediately followed with feedback and praise. When the student stopped playing, the experimenter provided feedback within 10 seconds. Feedback included first telling the student what he/she did well for each of the four dependent variables if without errors. For example, "You played all of the rhythms correctly and you did not hesitate at all." Immediately upon completion of the positive statement, the experimenter showed the student the marked up copy of the music and pointed to the markings while telling the student what errors were made. For example, "You played a wrong note here and here, you hesitated here and then repeated this measure." All feedback ended with a positive comment, such as "That was a good try! Let's try another one." The experimenter then saved this piece to be presented at a later time, presented the next of the fifteen pieces, and repeated the training procedure. If the student played a piece correctly, then the experimenter did not present this piece again and presented the next piece. The participants moved through the pieces on a rotating basis. Treatment continued until the student played all 15 training pieces correctly, without any errors.

Post-treatment. For post-treatment, the experimenter required students to play the post-treatment sight-reading pieces until data were stable. The procedure for post-treatment was identical to baseline. The experimenter provided general feedback within 10 seconds of completion (e.g., “Good try, you made a few mistakes, let’s try another one.”). She presented all new pieces that the students had not seen before and the students played each piece only once. She again collected data on musical errors made by the students while playing these pieces.

Inter-observer Agreement

The experimenter videotaped all sessions and later she and an independent observer viewed the tapes. Prior to data scoring, the experimenter and the observer conduct practice sessions until inter-observer agreement scores were 95% or better. The experimenter calculated inter-observer agreement by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. Both had a copy of all music to mark errors on while viewing the tapes. The experimenter defined agreement as both observers recording the occurrence of musical errors during each piece and disagreement as one observer recording the occurrence of a musical error and the other observer not recording an error. Inter-observer agreement was scored for 75% of randomly selected sessions. Agreement averaged 97% (range 93% to 99%), 99% (range 96% to 99%), and 99% (range 93% to 99%) for Students 1, 2, and 3 respectively.

Treatment Integrity

An independent observer also viewed the video tapes to score treatment integrity. Treatment integrity was scored for four randomly selected sessions per phase. For these sessions, the observer scored the teacher's implementation of the behavioral-skills training package. The observer marked off teacher behavior on checklists for baseline (see Appendix B), treatment (see Appendix C), and post-training (see Appendix B). The observer scored each item as either correct or incorrect, by placing a check mark next to each step that is completed correctly by the researcher. The observer calculated treatment integrity by dividing the number of teacher errors by the total number of correct plus incorrect items on the checklist and multiplying by 100%. Treatment integrity was 100% for all randomly selected sessions in each phase.

Social Validity

The experimenter assessed social validity with a survey presented to the participants at the end of the study. The experimenter developed the survey and included questions on their opinion of the importance of various sight-reading skills and their perceived improvement in sight-reading, etc. (see Appendix D). Furthermore, 10 music teachers who did not have regular contact with the subjects were asked to observe a tape of each student sight-reading during three baseline and three post-treatment sessions. The experimenter did not tell the teachers which sessions were which. She gave them a copy of each piece of music and asked them to score whether the students performed the piece correctly (see Appendix E). After observing and scoring all video clips, each music teacher scored the importance of music sight-reading for musicians on a scale of 1 to 10, with 10 being the highest.

Results

As shown in Figure 1, there was a systematic decrease in the percentage of note errors as treatment was implemented across students. For student 1, the average percentage of note errors during baseline was 11% (range 0% to 40%). The average percentage of note errors decreased to 1% (range 0% to 3%) during post-treatment. For student 2, the average percentage of note errors decreased from 24% (range 4% to 63%) during baseline to 6% (range 1% to 20%) during post-treatment. For Student 3, the average percentage of note errors during baseline was 5% (range 0% to 19%). During post-treatment, the average percentage of note errors decreased to 1% (range 0% to 6%).

Figure 1 *Percentage of Note Errors Played During Sight-Reading Across Sessions*

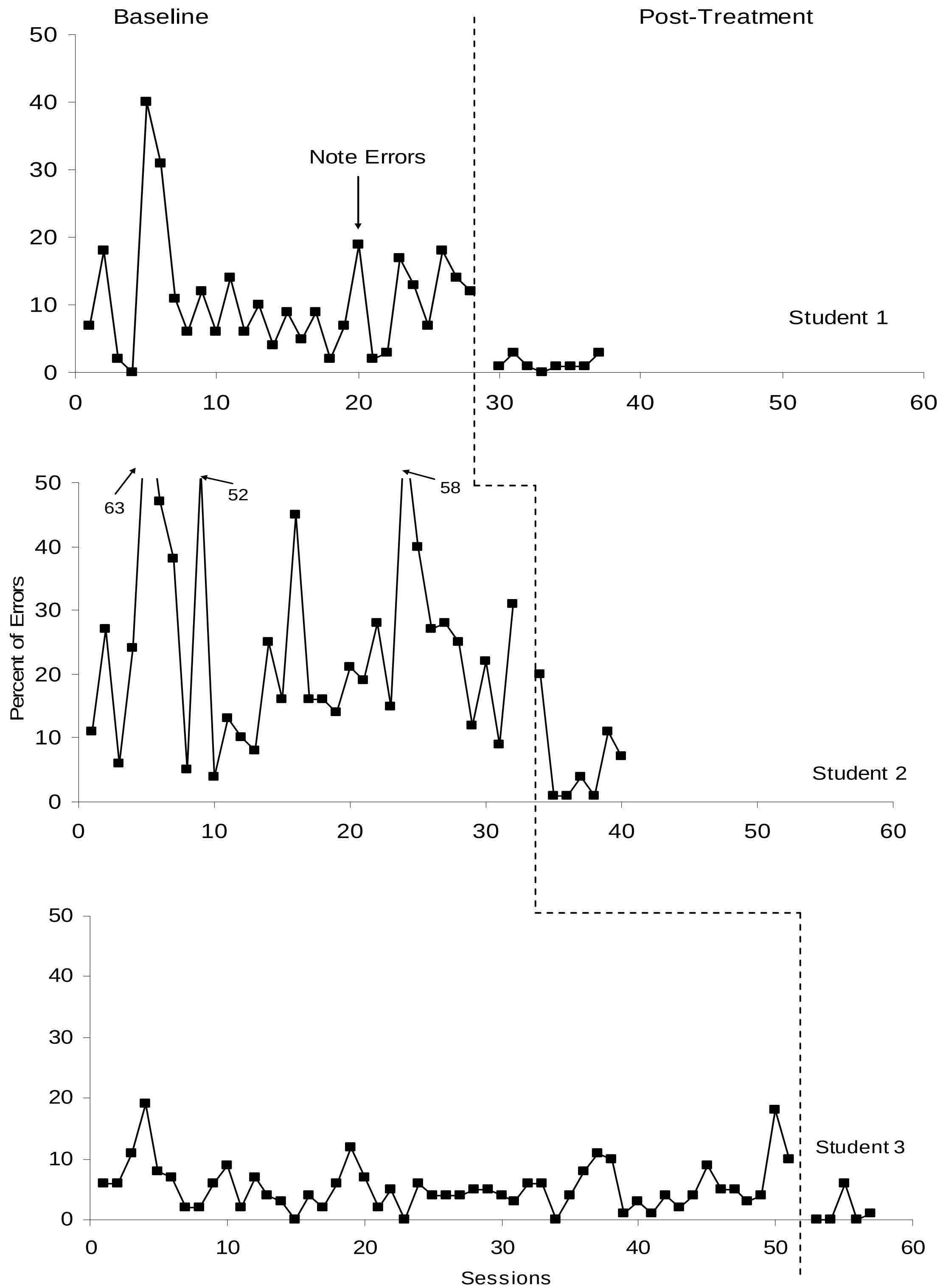


Figure 2 shows a systematic decrease in the percentage of rhythmic errors made during sight-reading as treatment was implemented across students. For student 1, the average percentage of rhythm errors during baseline was 42% (range 2% to 93%). The average percentage of rhythm errors decreased to 3% (range 0% to 9%) during post-treatment. For student 2, the average percentage of rhythm errors decreased from 59% (range 27% to 87%) during baseline to 7% (range 0% to 25%) during post-treatment. For Student 3, the average percentage of rhythm errors during baseline was 39% (range 1% to 83%). During post-treatment, the average percentage of rhythm errors decreased to 4% (range 0% to 9%).

Figure 2 *Percentage of Rhythm Errors Played During Sight-Reading Across Sessions*

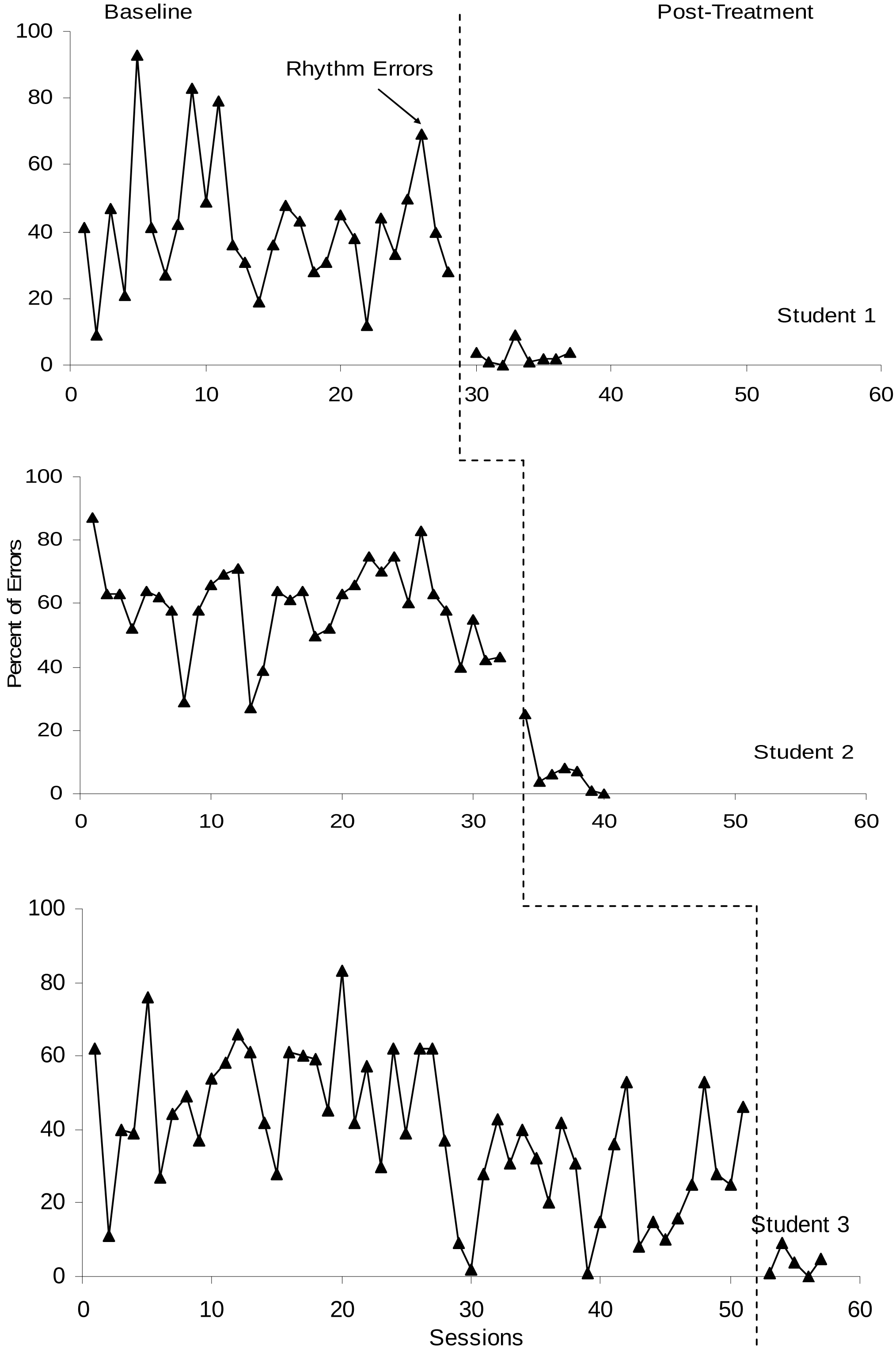
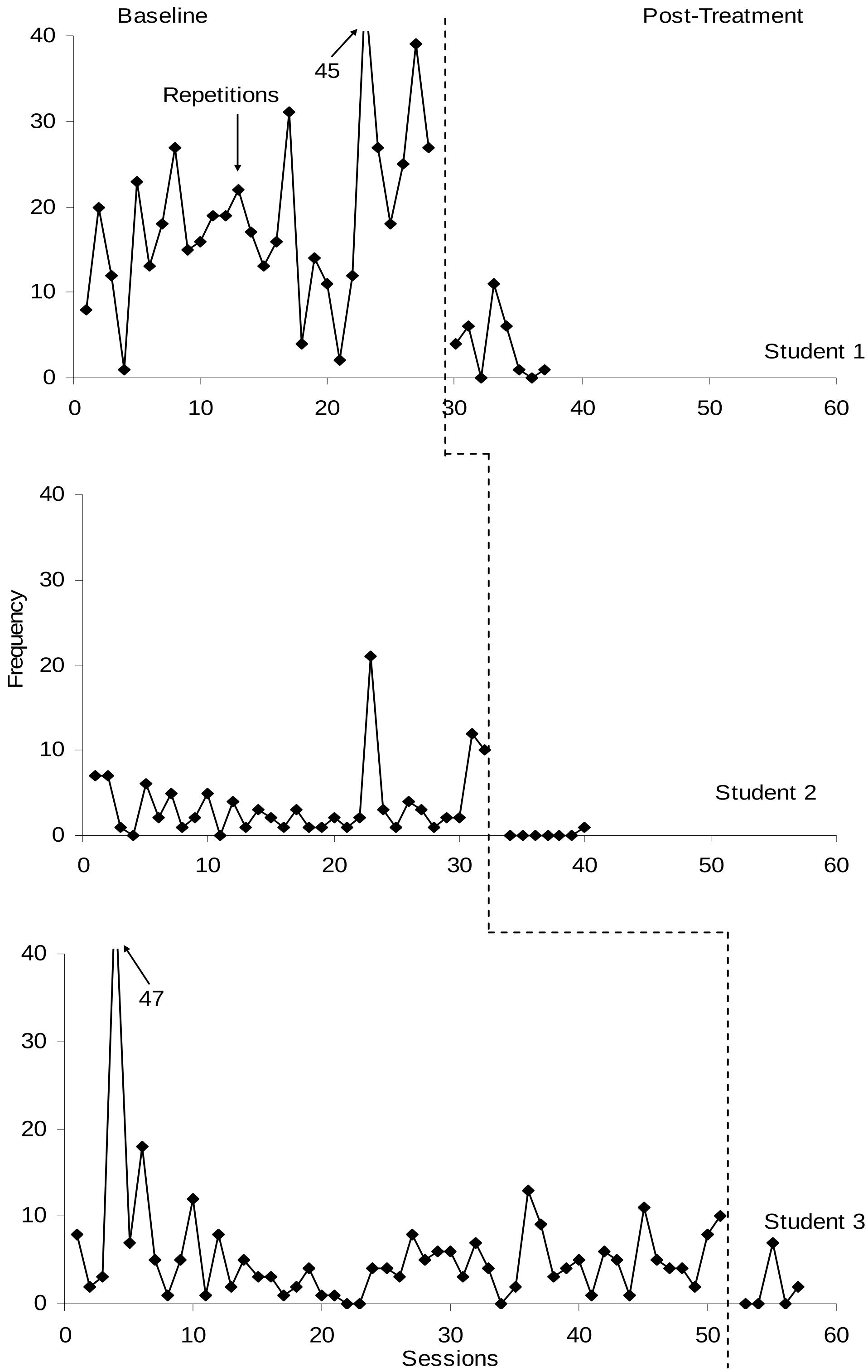


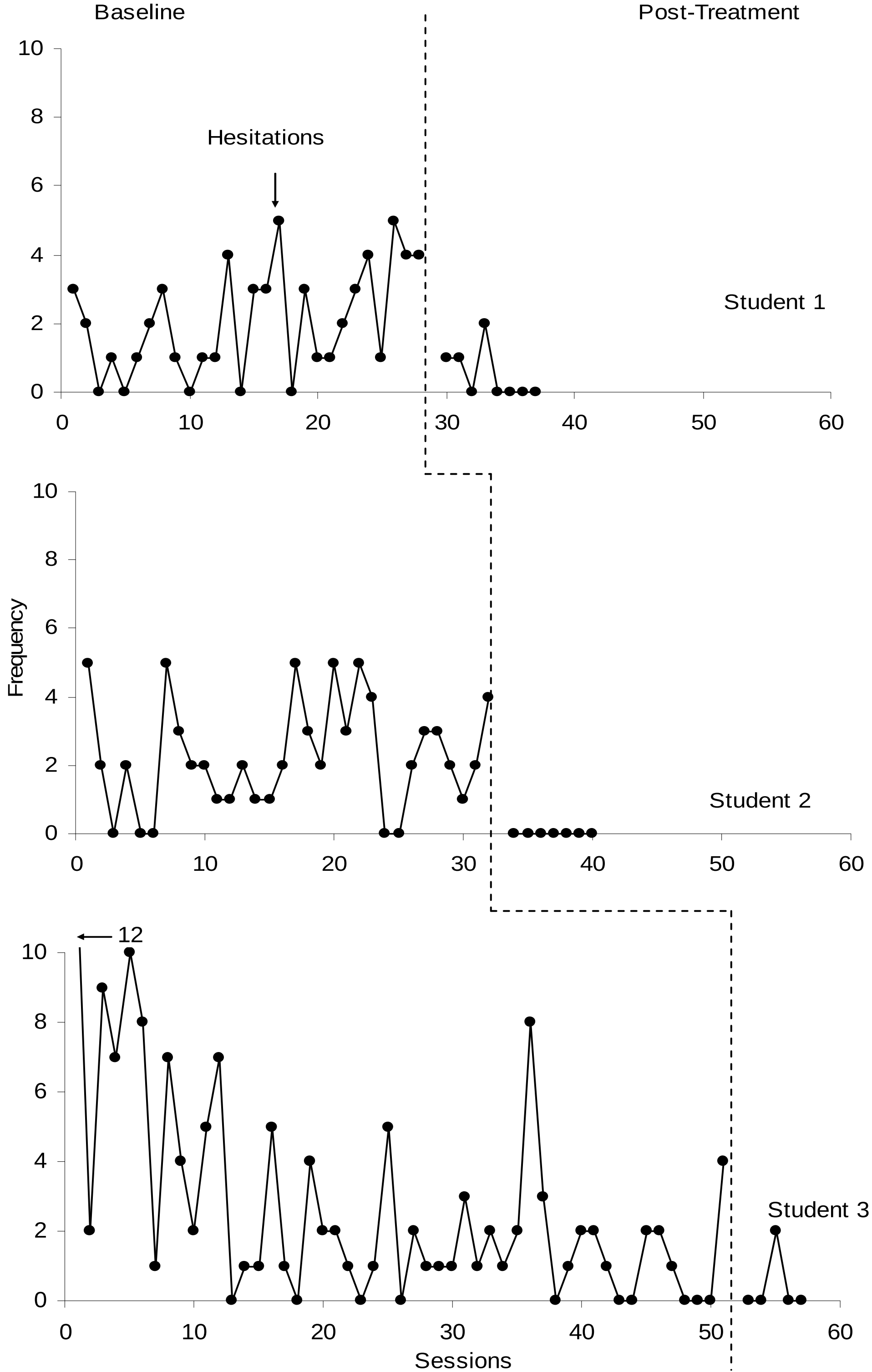
Figure 3 demonstrates a systematic decrease in the frequency of repetitions as treatment was implemented across subjects. For student 1, the average frequency of repetitions during baseline was 18 (range 1 to 45). The average frequency of repetitions decreased during post-treatment to 1 (range 0 to 11). For student 2, the average frequency of repetitions was 4 (range 0 to 21) during baseline and decreased to 0 (range 0 to 1) during post-treatment. For Student 3, the average frequency of repetitions during baseline was 6 (range 0 to 47). The average frequency of repetitions decreased to 2 (range 0 to 7) during post-treatment.

Figure 3 *Frequency of Repetitions Played During Sight-Reading Across Sessions*



As shown in Figure 4, there was a systematic decrease in the frequency hesitations as treatment was implemented across subjects. For student 1, the average frequency of hesitations was 2 (range 0 to 5) during baseline and decreased to 1 (range 0 to 2) during post-treatment. For student 2, the average frequency of hesitations during baseline was 2 (range 0 to 5) and decreased to 0 for all post-treatment sessions. For Student 3, the average frequency of hesitations during baseline was 3 (range 0 to 12). The average frequency of hesitations decreased to 0 (range 0 to 2) during post-treatment.

Figure 4 *Frequency of Hesitations Played During Sight-Reading Across Sessions*



Throughout all phases of the study, the students recorded the amount of time they spent practicing their instrument at home in a practice log. The minutes of practice time recorded in the logs did not increase with the presentation of treatment (see Appendix F). In fact, all three students recorded at least 2 weeks of little or no practice during treatment. Fluctuations in practice time remained somewhat consistent throughout the study. Therefore, the decrease in all dependent variables was not due to an increase in student practice time.

Student responses to the questionnaires about their opinion of the importance of various sight-reading skills demonstrated that the students believed the dependent variables were important skills for musicians. Student responses were “Agree” or “Strongly Agree” for all questions about the importance of each skill to musicians, as well as for the question about the effectiveness of the method used and the question about whether the improvement in their sight-reading was a worthwhile change. Furthermore, some students wrote comments about the study (i.e., “I believe the study was helpful.”; “I was able to correct my mistakes and learn an easier way to play through sight-reading.”; and “It helped with NYSSMA very much!”). Teacher responses demonstrated a clear difference between baseline and post-treatment sessions, with post-treatment sessions scored higher for all students. Questions asked for a rating of how well the student played for each of the dependent variables and a general question about how they sight-read overall. The average percentage of responses for baseline sessions was 61% “Disagree” or “Strongly Disagree” and 39% “Agree” or “Strongly Agree.” For post-treatment sessions, the average percentage of responses was 16% “Disagree” or “Strongly Disagree” and 84% “Agree” or “Strongly Agree” (see Table 1). The experimenter also asked all of the teachers how important sight-reading is to musicians on a scale of 1 to 10. All of the teachers responded between 8 and 10 for their perceived importance of sight-reading for musicians. Three teachers responded “8”, two responded “9”, and five responded “10.”

Table 1

Means of Teacher Responses to Videos of Students Sight-Reading

	Student 1		Student 2		Student 3	
	Baseline	Post-Treatment	Baseline	Post-Treatment	Baseline	Post-Treatment
Question						
1	1.1	2.3	1.3	2.6	1.1	2.1
2	0.6	1.8	1.5	2.3	0.9	1.4
3	0.8	2.4	1.8	2.7	1.4	2.4
4	0.6	2.1	1.7	2.4	1.1	2.0
5	1.4	2.3	1.8	2.6	1.4	2.2

Discussion

This study demonstrated the usefulness of a treatment package that includes general-case training, instructions, feedback, and rehearsal in improving sight-reading ability in advanced flute students. Because there was a systematic decrease in student incorrect sight-reading as the treatment was implemented across subjects, it can be concluded that the combination of behavioral-skills training and general-case training caused the change. This improvement of all three students' performance was not simply due to improvement with practice over time. All students recorded practice time throughout the study. The practice logs showed that practice time did not increase throughout the study for any of the students. Furthermore, none of the students met criteria for completing a piece until after the experimenter implemented treatment. Student 3 remained in baseline for the greatest number of sessions and although this student demonstrated some improvement during the extended baseline, the change after treatment was much more dramatic. Furthermore, all of these students previously took private lessons and participated in NYSSMA for at least five years. All students had the same private flute teacher, the experimenter, who trained them on sight-reading using typical teaching methods. Even with years of previous training, the training procedure greatly improved the students' sight-reading. This suggests that the strategies used for the selection of training exemplars may be the key aspect of the treatment package that made a difference compared to regular sight reading practice.

The results of the current study are an important contribution to the literature because they impact two bodies of research, that of behavioral-skills training and general-case training. The results of this study are consistent with the findings of previous studies in that there is a reliable demonstration of the usefulness of a behavioral skills training package for teaching music-related behavior. More specifically, the current findings demonstrate the effectiveness of behavioral skills training for teaching music behavior (Dib & Sturmey, 2007; Eisenstein, 1976; Salzberg & Salzberg, 1981). There are some differences between previous and current studies. The current study is the only behavioral study on music behavior which investigated sight-reading skills. Therefore, the current study is important because it demonstrates a novel use for two training procedures, general-case training and behavioral skills training. This study also adds to and expands upon the body of research investigating general-case training for the promotion of response generalization (Ducharme & Feldman, 1992; Horner & McDonald, 1982; Hupp & Mervis, 1981; Neef, et al., 1990; Sprague & Horner, 1984). Although music sight-reading is a skill of generalization from learned music to new music, no research so far has investigated the use of general-case training for music education. Only Dib and Sturmey (2007) addressed stimulus generalization and no studies addressed response generalization.

Another contribution of the current study was the creation of a novel 2-dimention matrix to sample the "instructional universe" for music key and time signatures. The procedure used to identify baseline, treatment, and post-treatment stimuli (music pieces) demonstrated the systematic creation of stimuli described in the general-case training research. Previous generalization research attempted to make a distinction between multiple-exemplar training and general-case training. In reviewing the literature, this author determined that the difference between the two was in the way training stimuli were created and/or chosen. Both procedures utilize the presentation of more than one exemplar. Therefore, both are examples of "multiple-exemplar" training. It seems that rather than being a completely different technique, general-case training may actually be a more specified and systematic way of selecting training stimuli for multiple-exemplar training. General-case training may be described as general-case training "done well." If this is the case, then studies comparing multiple-exemplar training and general-case training are actually demonstrations of how to create and/or choose training stimuli to produce the most generalization during testing.

To create training stimuli that support the most generalization, attention should be paid to the quality, rather than the quantity, of exemplars selected. The studies of Hupp and Mervis (1981) and Horner and McDonald (1982) support the argument that it is not the number of exemplars used, but rather how representative the exemplars are that is important in the programming of generalization. In fact, fewer exemplars may be more effective than many exemplars if only "good" (representative) exemplars are used (Hupp & Mervis, 1981). The current study supports this literature in that the experimenter used only 15 training pieces to improve sight-reading for all three students. The experimenter specifically created and chose these training pieces to sample all of the possible time- and key-signatures that the students would come into contact with.

As mentioned previously, the private teacher provided these students with sight-reading training for at least 5 years. Typical teaching methods for promoting sight-reading for flute students are comparable to multiple-exemplar training. Teachers choose pieces or sections of pieces that are unfamiliar to the student and present these to the student to play without having practiced. Teachers present students with various musical pieces, which should sample various key- and time-signatures; however, the teachers probably do not select all of the possibilities in the "instructional universe." For example, flute students who play in the school band will learn flat notes before sharp notes because the majority of band music contains flats in the flute parts. Furthermore, method books for flute typically begin with natural, then flat notes. Students may not learn sharps for months and even after learning some flats and sharps, students may not see a piece with more than 3 or 4 flats or sharps until about a year after beginning lessons. Although this order of teaching musical notes to flute players may be effective in teaching music reading and playing, results of the current study suggest that selecting specific training pieces to sample time- and key-signatures would increase generalization to new music. The way in which the experimenter selected training exemplars may be the key aspect of the treatment package that made a difference compared to regular sight-reading training.

A limitation to the current study is that it looks only at sight-reading rather than a broader range of student flute-playing behavior, such as tone, technique, and improvisation. Future research should investigate the use of general-case training and/or behavioral skills training in other flute-playing behavior, as well as in the teaching of other instruments. Another limitation is that, although the combination of general-case and behavioral skills training was effective, the current study did not allow for the separation of these two procedures in order to determine if one was more effective than the other or if one of the components would have been effective on its own. For example, it is unclear whether behavioral skills training would have been effective without general-case training. Although research has found that behavioral skills training may be a robust package to change music behavior, research has not yet identified the effective components of this package. Research could investigate the components of the package individually to determine if they would be as effective separately.

In summary, the current study was effective in decreasing musical errors made by advanced flute students while sight-reading music. There was a dramatic improvement in all three student’s sight-reading even though they were all previously trained on sight-reading for a minimum of 5 years. Therefore, it can be concluded that the training procedure was more effective than typical teaching methods. The development of the matrix enabled the experimenter to identify training pieces that sampled the instructional universe. The experimenter selected the fewest training exemplars required to sample all of the key- and time-signatures. The effectiveness of the use of these 15 training pieces suggests that general-case training should be used to teach music sight-reading. Future research should investigate other teaching situations in which general-case training would be more effective than typical methods for increasing generalization of responses.

Appendix A

Sight-Reading Stimuli Matrix

Flats/Sharps in the Key Signature

Time Signatures	None	Bb	Bb, Eb	Bb, Eb, Ab	Bb, Eb, Ab, Db	Bb, Eb, Ab, Db, Gb	Bb, Eb, Ab, Db, Gb, Cb	Bb, Eb, Ab, Db, Gb, Cb, Fb	F#	F#, C#	F#, C#, G#	F#, C#, G#, D#	F#, C#, G#, D#, A#	F#, C#, G#, D#, A#, E#	F#, C#, G#, D#, A#, E#, B#
4/4	T 1	B/P 8	B/P 15	B/P 22	B/P 29	B/P 36	B/P 43	B/P 50	B/P 57	B/P 64	B/P 71	B/P 78	B/P 85	B/P 92	T 15
6/8	B/P 1	T 2	B/P 16	B/P 23	B/P 30	B/P 37	B/P 44	B/P 51	B/P 58	B/P 65	B/P 72	B/P 79	B/P 86	T 14	B/P 99
2/2	B/P 2	B/P 9	T 3	B/P 24	B/P 31	B/P 38	B/P 45	B/P 52	B/P 59	B/P 66	B/P 73	B/P 80	T 13	B/P 93	B/P 100
12/8	B/P 3	B/P 10	B/P 17	T 4	B/P 32	B/P 39	B/P 46	B/P 53	B/P 60	B/P 67	B/P 74	T 12	B/P 87	B/P 94	B/P 101
3/8	B/P 4	B/P 11	B/P 18	B/P 25	T 5	B/P 40	B/P 47	B/P 54	B/P 61	B/P 68	T 11	B/P 81	B/P 88	B/P 95	B/P 102
2/4	B/P 5	B/P 12	B/P 19	B/P 26	B/P 33	T 6	B/P 48	B/P 55	B/P 62	T 10	B/P 75	B/P 82	B/P 89	B/P 96	B/P 103
9/8	B/P 6	B/P 13	B/P 20	B/P 27	B/P 34	B/P 41	T 7	B/P 56	T 9	B/P 69	B/P 76	B/P 83	B/P 90	B/P 97	B/P 104
3/4	B/P 7	B/P 14	B/P 21	B/P 28	B/P 35	B/P 42	B/P 49	T 8	B/P 63	B/P 70	B/P 77	B/P 84	B/P 91	B/P 98	B/P 105

T = Treatment

B/P = Baseline/Post-treatment

Appendix B

Treatment Integrity Checklist

Directions: Please mark a "☑" on the line to the left of the statement if the researcher completed the step correctly. Leave the line blank if the step was not completed correctly.

Baseline/Post-Treatment

___ Researcher presents student with line of music to sight-read

___ Researcher does not provide instructions

___ Researcher marks errors on her copy while the student plays

___ Researcher provides general feedback within 10 seconds

Appendix C

Treatment Integrity Checklist

Directions: Please mark a "☑" on the line to the left of the statement if the researcher completed the step correctly. Leave the line blank if the step was not completed correctly.

Treatment

___ Researcher presents student with line of music to sight-read

___ Researcher provides instructions

    ___ Reminds student to look at key signature

    ___ Reminds student to look at time signature

    ___ Reminds student to look for accidentals

___ Researcher marks errors on her copy while the student plays

___ Researcher provides feedback within 10 seconds

___ Researcher shows student her marked up music sheet

    ___ Positive Feedback:

        ___ Notes ___ Rhythms ___ Hesitations ___ Repetitions ___ Frustrated Behavior

    ___ Corrections:

        ___ Notes ___ Rhythms ___ Hesitations ___ Repetitions ___ Frustrated Behavior

___ Researcher gives student new piece to play (and repeats all steps)

Appendix D

Subject Questionnaire

Please indicate the extent to which you agree with each of the following statements regarding the study on music sight-reading.

Rating Key:
SD = Strongly Disagree **D** = Disagree **A** = Agree **SA** = Strongly Agree

	SD	**D**	**A**	**SA**
1. Musicians should try to reduce note errors (defined as playing a pitch that is different from what is written in the music).				
2. Musicians should try to reduce rhythmic errors (defined as playing a note or rest longer or shorter than its written value).				
3. Musicians should try to reduce repetitions (defined as re-playing a part or all of the measure instead of moving on to the next measure).				
4. Musicians should try to reduce hesitations (defined as any pause that interrupts playing through the measure).				
5. The method used to improve my sight-reading was a good one.				
6. The improvement in my sight-reading was a worthwhile change.				

Additional Comments:

____

____

____

____

____

Appendix E

Music Teacher Observation Survey

Please indicate the extent to which you agree with each of the following statements regarding the student you observed sight-reading in the video tape.

Rating Key:
SD = Strongly Disagree **D** = Disagree **A** = Agree **SA** = Strongly Agree

	SD	**D**	**A**	**SA**
1. The student made few note errors (defined as playing a pitch that is different from what is written in the music).				
2. The student made few rhythmic errors (defined as playing a note or rest longer or shorter than its written value).				
3. The student made few repetitions (defined as re-playing a part or all of the measure instead of moving on to the next measure).				
4. The student made few hesitations (defined as any pause that interrupts playing through the measure).				
5. Overall, the student sight-read well.				

Additional Comments:

____

____

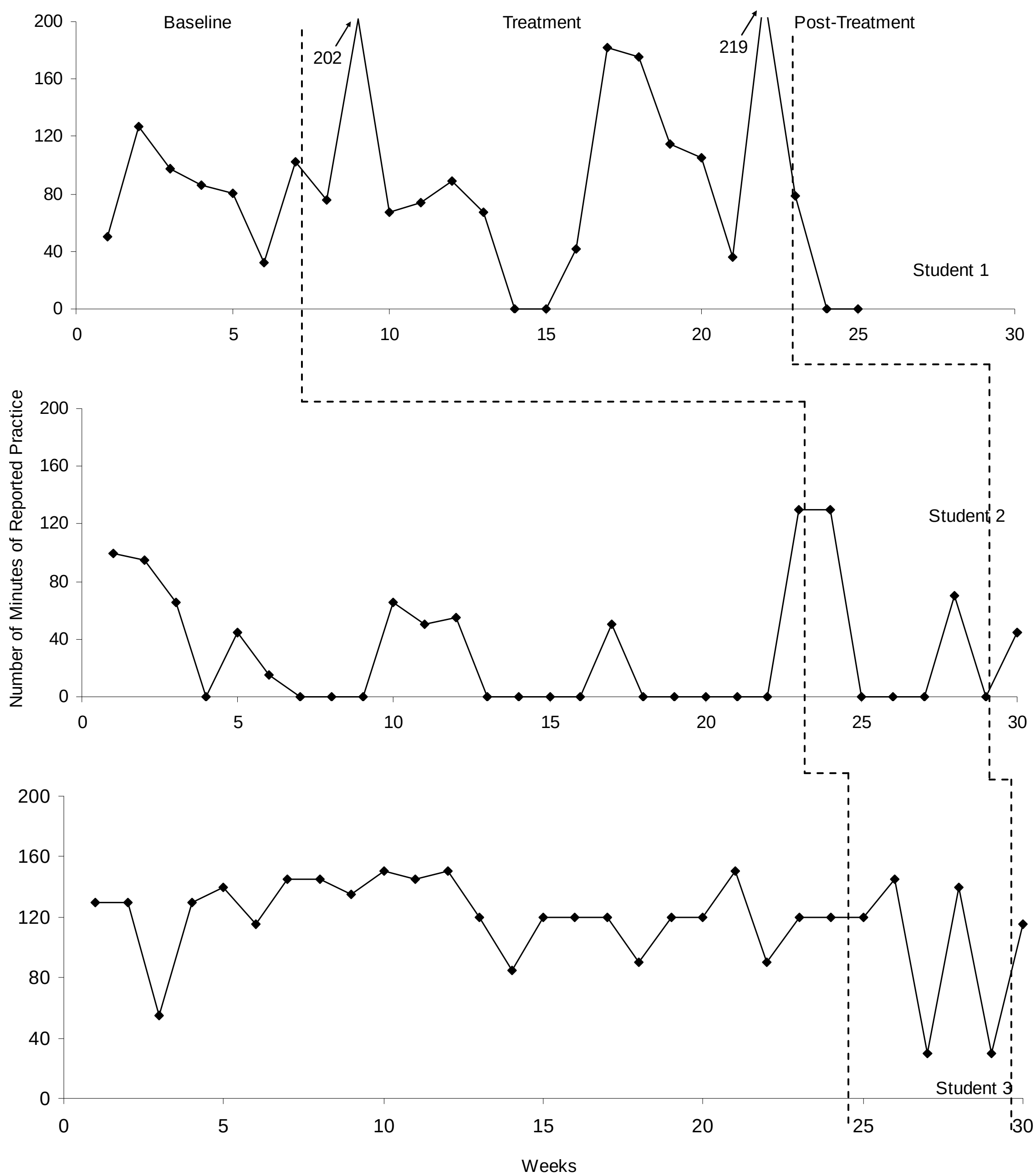
____

____

____

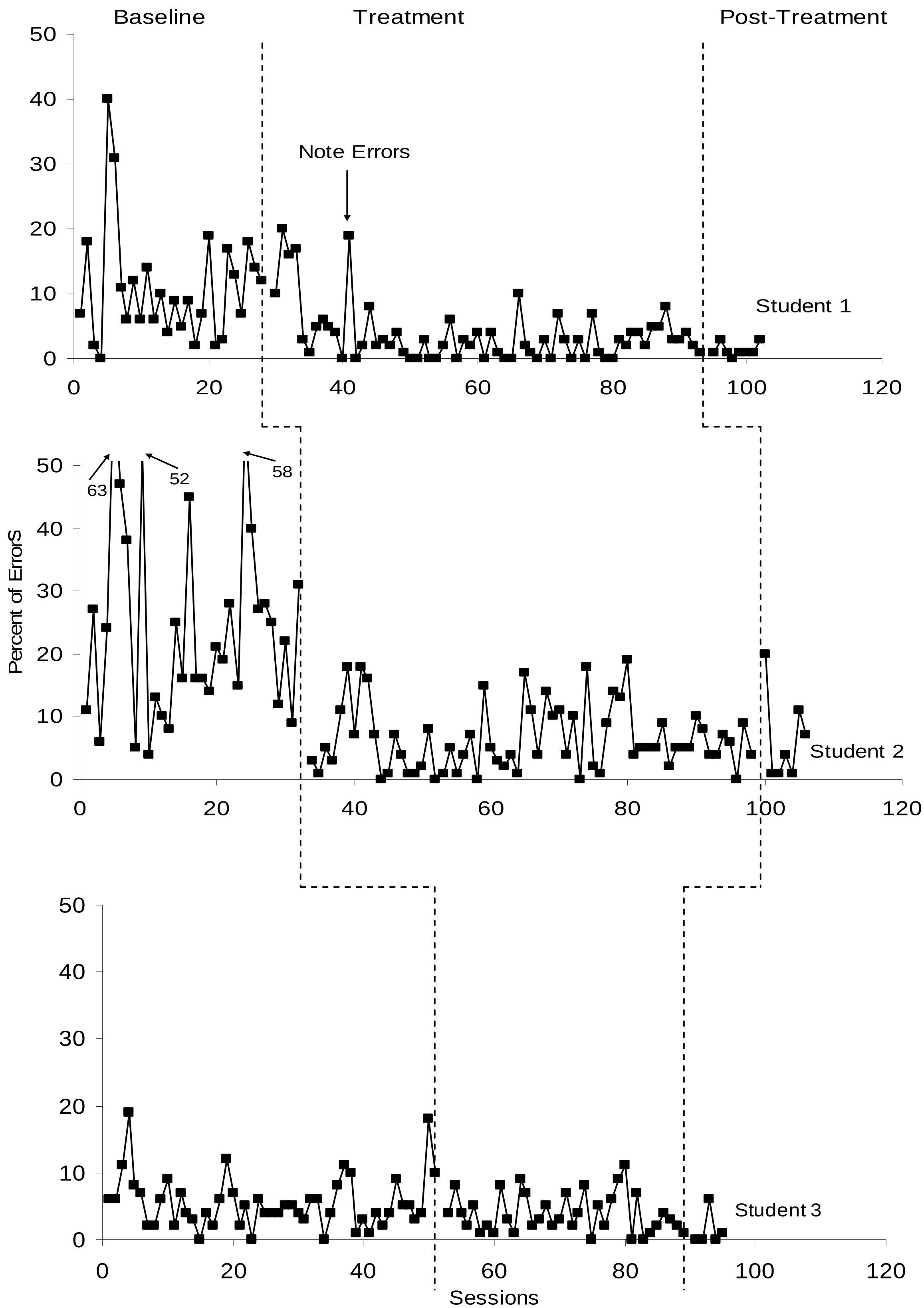
Appendix F

Graph of Practice Time Recorded in Student Logs

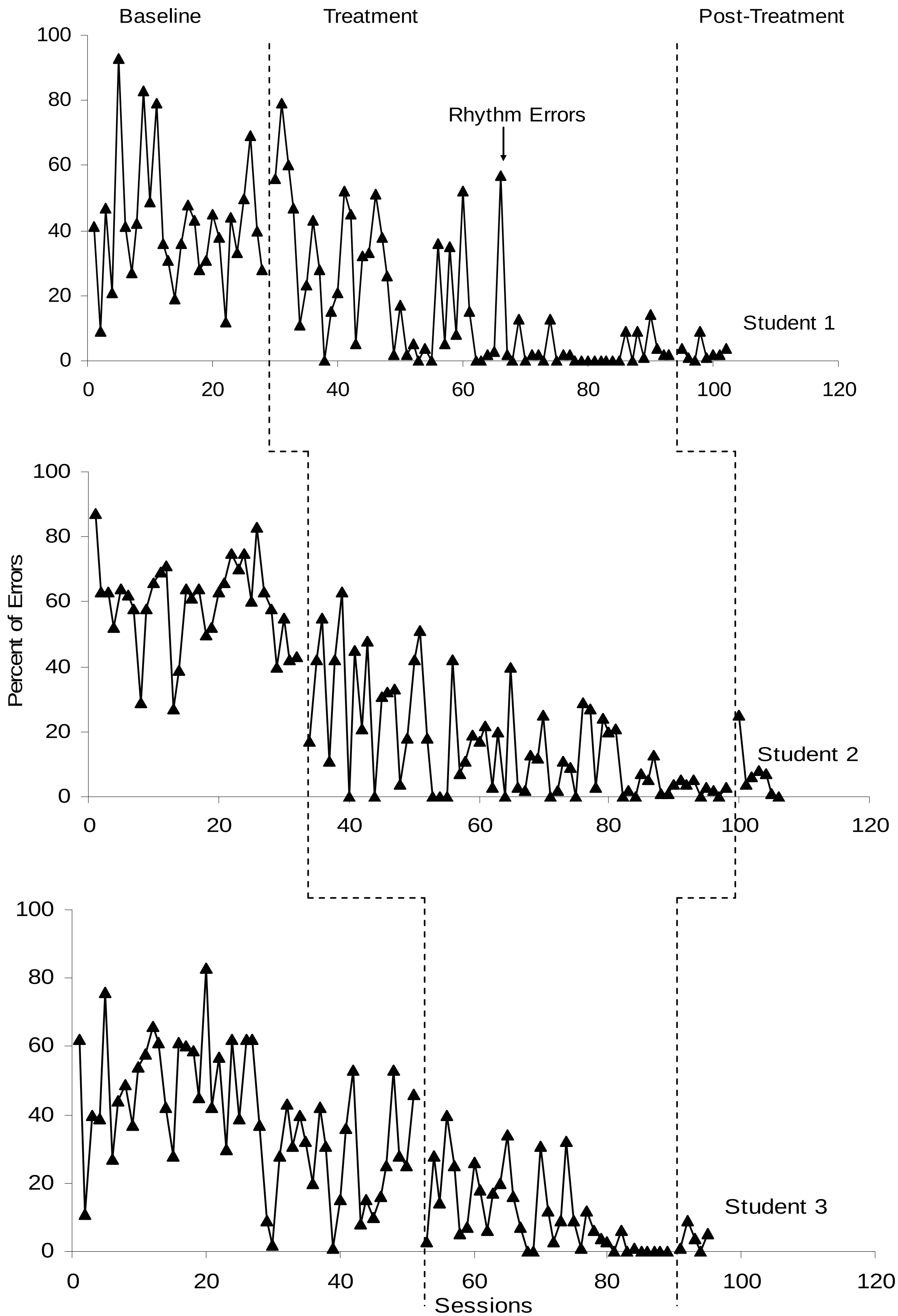


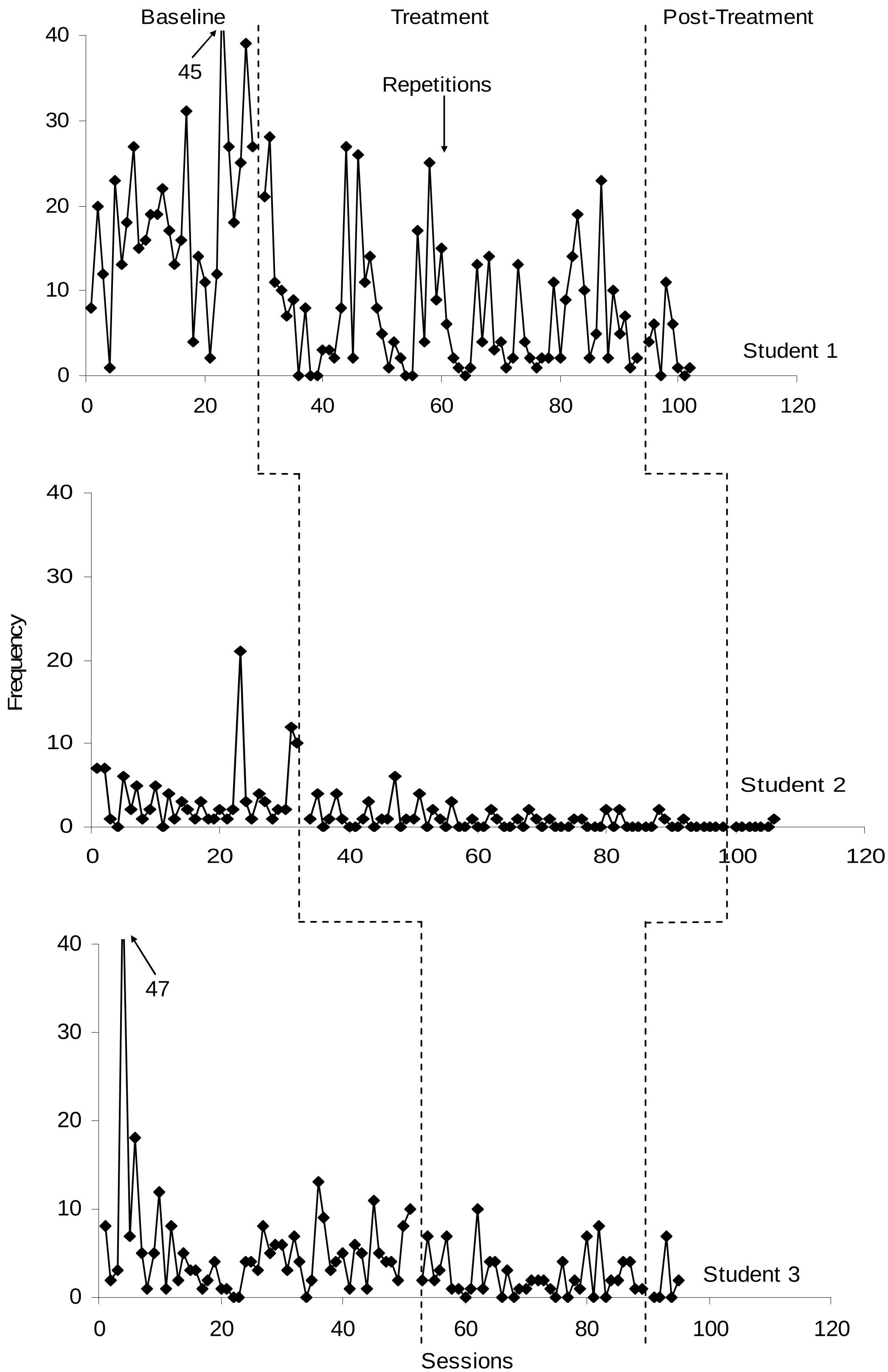
Appendix G

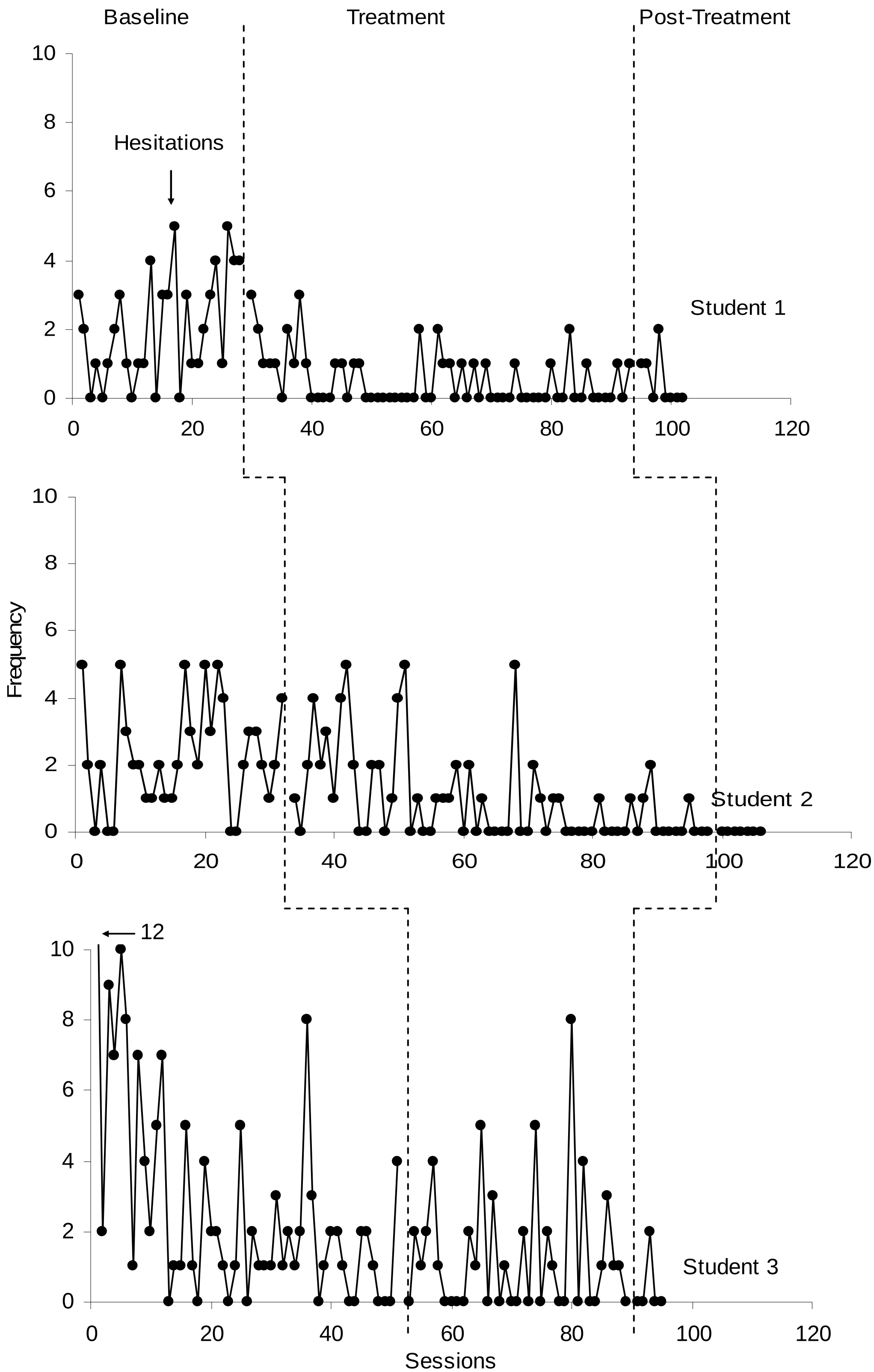
Graphs Including Treatment Phase



Baseline
Treatment
Post-Treatment
Note Errors
Student 1
63
52
58
Student 2
Percent of ErrorS
Student 3
Sessions
0
10
20
30
40
50
0
20
40
60
80
100
120


Baseline
Treatment
Post-Treatment
Rhythm Errors
Student 1
Student 2
Student 3
Percent of Errors
Sessions
0
20
40
60
80
100
120


Baseline
Treatment
Post-Treatment
45
Repetitions
Student 1
Student 2
47
Student 3
Frequency
Sessions
0
20
40
60
80
100
120
10
30



Baseline
Treatment
Post-Treatment
Hesitations
Student 1
Student 2
Student 3
12
Frequency
Sessions
0
2
4
6
8
10
0
20
40
60
80
100
120

Bibliography

Berens, N. M. & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis, 40*, 45-71.

Catania, A. C. (1992). *Learning* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Colvin, G. T. & Horner, R. H. (1983). Experimental analysis of generalization: An evaluation of a general case program for teaching motor skills to severely handicapped learners. In D. Hogg & P. Mittler (Eds.). *Aspects of competence in severely mentally handicapped people* (Vol. 11). Sussex, England: John Wiley & Sons Limited.

Dib, N. E. (2007). *The Acquisition, Modification, and Maintenance of Music-Related Behavior: A Review and Implications for Future Research*. Unpublished master's Thesis, Queens College, New York.

Dib, N.E., & Sturmey, P. (2007). The effects of verbal instruction, modeling, rehearsal, and feedback on correct posture during flute playing. *Behavior Modification, 31*, 382-388.

Ducharme, J. M. & Feldman, M. A. (1992). Comparison of staff training strategies to promote generalized teaching skills. *Journal of Applied Behavior Analysis, 25*, 165-179.

Eisenstein, S. R. (1976). A successive approximation procedure for learning music symbol names. *Journal of Music Therapy, 13*, 173-179.

Fields, L., Matneja, P., Varelas, A., Belanich, J., Fitzer, A., & Shamoun, K. (2002). The formation of link perceptual classes. *Journal of the Experimental Analysis of Behavior, 78*, 271-290.

Horner, R. H. & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *Journal of the Association for the Severely Handicapped, 8,* 7-20.

Hughes, C., Harmer, M. L., Killian, D. J., & Niarhos, F. (1995). The effects of multiple exemplar self-instruction training on high school students' generalized conversational interactions. *Journal of Applied Behavior Analysis, 28,* 201-218.

Hupp, S. C. & Mervis, C. B. (1981). Development of generalized concepts by severely handicapped students. *Journal of the Association for the Severely Handicapped, 6,* 14-21.

Luciano, C., Becerra, I. G., & Valverde, M. R. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of Applied Behavior Analysis, 87,* 349-365.

Neef, N. A., Lensbower, J., Hockersmith, I., DePalma, V., & Gray, K. (1990). In vivo versus simulation training: An interactional analysis of range and type of training exemplars. *Journal of Applied Behavior Analysis, 23,* 447-458.

Poulson, C. L. & Nunes, L. R. P. (1988). The infant vocal-conditioning literature: A theoretical and methodological critique. *Journal of Experimental Child Psychology, 46,* 438-450.

Reeve, S. A., Reeve, K. F., Buffington Townsend, D., & Poulson, C. L. (2007). Establishing a generalized repertoire of helping behavior in children with autism. *Journal of Applied Behavior Analysis, 40,* 123-136.

Salzberg, R. S. & Greenwald, M. A. (1977). Effects of a token system on attentiveness and punctuality in 2 string instrument classes. *Journal of Music Therapy, 14,* 27-38.

Salzberg, R. S. & Salzberg, C. L. (1981). Praise and corrective feedback in the remediation of incorrect left-hand positions of elementary string players. *Journal of Research in Music Education, 29*, 125-133.

Stokes, T. F. & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Sprague, J. R. & Horner, R. H. (1984). The effects of single instance, multiple instance, and general case training on generalized vending machine use by moderately and severely handicapped students. *Journal of Applied Behavior Analysis, 17*, 273-278.

Wolpow, R. I. (1976). The independent effects of contingent social and academic approval upon the music on-task and performance behaviors of profoundlyretarded adults. *Journal of Music Therapy, 13*, 29-38.