12-2018

Leveraging Stimulus Equivalence to Teach Piano to Children with Autism

Krystin Katherine Hussain

Leveraging Stimulus Equivalence to Teach Piano to Children with Autism

by

Krystin Katherine Hussain

A thesis submitted to the School of Behavior Analysis of
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Applied Behavior Analysis and Organizational Behavior Management

Melbourne, Florida
December, 2018

We the undersigned committee hereby approve the attached thesis, "Leveraging Stimulus Equivalence to Teach Piano to Children with Autism" by Krystin Katherine Hussain.

Catherine A. Nicholson, Ph.D.
Assistant Professor
School of Behavior Analysis
Committee Chair

Ivy Chong, Ph.D.
Senior Vice President
Home, School, and Center- based Services, May Institute

Heidi H. Edwards, Ph.D.
Professor and Associate Head
School of Arts and Communication

Jose A. Martinez-Diaz, Ph.D.
Professor and Head
School of Behavior Analysis

Abstract

Title: Leveraging Stimulus Equivalence to Teach Piano to Children with Autism

Author: Krystin Katherine Hussain

Advisor: Catherine A. Nicholson, Ph. D.

Music-based interventions have been shown to benefit individuals diagnosed with autism spectrum disorder (ASD) by improving deficits such as social behaviors, communication, and vocalizations, as well as reducing behavioral excesses such as stereotypies (Hill, 2015). The purpose of the current study was to evaluate the effects of equivalence-based instruction (EBI) on acquisition of piano skills, novel piano performance, and generalization and maintenance of taught and untaught piano skills among children with autism. Training consisted of auditory-visual musical stimuli in a matching-to-sample format. Training was conducted using simultaneous matching in a one-to-many arrangement (relations between one stimulus are trained to multiple others). Learners selected letters, music notation, and piano keys when given an auditory stimulus. Following training, post-tests were conducted to test the emergence of novel untrained relations and generalization. Maintenance probes were conducted at least one week following the final post-test. All participants demonstrated novel piano skills and scored high on maintenance probes. Results suggest the efficacy of EBI in teaching and maintaining piano skills.

Table of Contents

List of Figures

List of Tables

Acknowledgement

I would first like to thank my thesis advisor Dr. Nicholson. Your guidance, support, and patience were fundamental in the success of this project. I would like to thank you for helping me to grow as a writer, researcher, and person.

To Dr. Chong and Dr. Edwards, thank you for your support and thoughtful recommendations throughout this project.

To Michael, thank you for helping me when I needed it most and guiding me forward when I felt like giving up. Your constant encouragement, helpful feedback, and endless kindness are more appreciated than I can ever express.

To Jeanine, Kirstin, Kayla, and Cherie, thank you for your support on this project. Your feedback and guidance will always be appreciated.

To Dr. Harvey, thank you for introducing me to behavior analysis and helping me to find a field that I love.

To my friends and family for always making me feel loved and cherished. I will always appreciate knowing you believe in me and continue to cheer me on.

Dedication

I dedicate this thesis to my parents, Faud and Sharon Hussain. Thank you for your continued support throughout this journey.

Introduction

Autism spectrum disorder (ASD) is a developmental disability, typically characterized by behavioral excesses classified as “unnatural” or “inappropriate,” such as stereotyped and/or repetitive use of objects, motor movements, and/or speech (Cunningham & Schreibman, 2009). Similar stereotypic behavior can be observed in neurotypically developing infants and toddlers, though it diminishes over time. In contrast, stereotypies tend to persist in individuals with developmental disabilities (Motor Disorders, 2017). Additionally, stereotypy can interfere with learning by competing with more desirable behaviors, preventing individuals from acquiring new skills (Koegel & Covert, 1972; Lovaas, Litrownik, & Mann, 1971).

Those with ASD also present with several deficits, such as disinterest in or inability to engage in leisure, social and communicative behaviors. These deficits can hinder a child’s future success because many of the behavioral deficits – social-emotional reciprocity, eye contact, understanding of gestures, integration of verbal and nonverbal communication – can lead to difficulty developing relationships with other people (American Psychiatric Association, 2013). In addition, deficits in communication and social skills often prevent them from engaging in appropriate leisure activities (Blum-Dimaya et al., 2010), which would compete with or even replace stereotypic and other undesirable behaviors. There hasn't been as much emphasis on leisure in the literature as on language and problem behavior. However, some investigators have taught skills such as shaping and playing with clay (Vuran, 2008), using activity schedules on the iPod touch (Carlile et al., 2013), playing Guitar Hero© (Blum-Dimaya et al., 2010), and using video prompting to teach aquatic play skills (Yanardag, Akmanoglu, & Yilmaz, 2013).

Individuals with ASD may also enjoy playing a musical instrument, yet this skill may be overlooked due to prioritizing other areas, such as language development. Still, autism interventionists may want to consider including music education when they develop individualized curricula for their learners. Comparisons between musicians and non-musicians in the general population have revealed that musical training results in better verbal, mathematical, and visual-spatial performance (Schlaug, Norton, Overy, & Winner, 2005). These benefits could be due to musicians' history with transforming a visual stimulus (symbols on sheet music) into motor movement (playing the correct notes) while receiving auditory feedback (listening to the played notes), which is a fairly complex set of behaviors to be engaging in simultaneously. Potentially, practicing these skills may lead to development of other auditory, motor, and multimodal response classes (Schlaug et al., 2005; Moreno & Farzan, 2015). For children with ASD, not only can instruction in music teach an appropriate leisure skill, it can also help improve fine motor movement and auditory discrimination, replace undesirable behavioral excesses, and even engender social interactions (Whipple, 2004).

Whipple (2004) conducted a meta-analysis of experimental studies comparing musical intervention to no musical intervention for children with ASD residing in the United States. Dependent variables in the studies included social behaviors (e.g., attention to a task, self-stimulation), communication (e.g., vocalizations, eye contact, speech/sign), and cognitive skills (e.g., vocabulary, compliance with motor tasks, academic tasks). Independent variables including social stories set to music, sung instructions, picture identification, language-based songs, music as reinforcement for other tasks, and musical accompaniment in activities. The studies utilized one of three approaches to music intervention: discrete trial instruction (i.e., instructor-led with systematic presentation of trials, prompting correct responding, providing positive reinforcement, and error correction; Leblanc, Ricciardi, & Luiselli, 2005), developmental socialpragmatic application (i.e., following the child's lead in a more naturalistic approach; Ingersoll, 2010), and "contemporary ABA" (i.e., a combination of the first two in a naturalistic, modestly structured environment; Whipple, 2004). Whipple found that music intervention resulted in an increase in social behaviors, communication, body awareness, coordination of vocalizations, comprehension of vocabulary, and attention. In addition, decreases in stereotypy, self-stimulation, inappropriate social behaviors, and anxiety were observed among the participants. Subsequent studies have found similar benefits such as increased social behaviors (Eren, 2015), speech production (Lim & Draper, 2011), and joint attention (Kim, Wigram, & Gold, 2008), as well as decreased stereotypy (Lanovaz, Sladeczek, & Rapp, 2011). Although Whipple found a positive direction for all calculated effect sizes, her review of the literature highlighted the need for more research on music instruction, particularly on how to teach children with ASD to play an instrument and the long-term benefits of doing so.

Teaching Learners With ASD

Various treatments are currently implemented to treat ASD; however, the evidence-based procedures derived from the science of applied behavior analysis (ABA) are recognized as the most effective. Applied behavior analysts study the relationship between behavior and the environment to solve problems of social significance with the goal of creating lasting behavior change (Cooper, Heron & Heward, 2007). Previous literature has illustrated that the implementation of ABA interventions leads to medium to large positive effects in areas such as language development, communication, intellectual functioning, and adaptive behavior (Makrygianni, Gena, Katoudi, & Galanis, 2018; Virués-Ortega, 2010). ABA focuses on changing *behaviors*, such as reducing problem behavior (e.g., stereotypies, aggression, tantrums; Martens, Daly, & Ardoin, 2015) and teaching and increasing communication (e.g., labeling, requesting; Ogletree & Oren, 2001), social skills (Mayer, Sulzer-Azaroff & Wallace, 2014) appropriate play and leisure activities (Blum-Dimaya, Reeve, Reeve, & Hoch, 2010; Carlile, Reeve, Reeve, & DeBar, 2013; Vuran, 2008), and generativity (i.e., recombining mastered skills to respond appropriately to novel situations; Grey & Hastings, 2005).

Playing music is an example of generative behavior; that is, once a minimum number of skills are directly taught, the learner can exhibit those skills under a wide array of stimulus conditions, and blend and recombine those skills to solve novel problems (Johnson, 2015). For example, once a student has learned to read music notation and execute certain fine motor behaviors directed to an instrument, the learner can play any song without having to memorize it, so long as there is sheet music to play from. Generativity is a key outcome of learning to play music (Cross et al., 2013) and interestingly, research in applied behavior analysis has identified several procedures designed to facilitate generative behavior, one of them being equivalence-based instruction.

Equivalence-Based Instruction

Equivalence-based instruction (EBI) is a teaching arrangement that leverages a kind of learning that occurs without direct teaching; specifically, learners can derive relationships between different stimuli. This behavioral phenomenon is called *stimulus equivalence.* Stimulus equivalence is defined as the emergence of correct responding to stimulus-stimulus relations that are untrained (not reinforced) subsequent to the reinforcement of *other* stimulus-stimulus relations (Cooper, Heron, & Heward, 2007). Stimulus equivalence is rooted in mathematical theory: If A=B and B=C, then A=C. Capital letters are typically used to denote different types of instructional material and combinations of the letters denote relationships between stimuli (Pytte & Fienup, 2012). For example, a picture of a dog is designated as stimulus A; the spoken word “dog” is designated stimulus B, and the written word “DOG” is designated stimulus C. If a learner is taught that a picture of a dog (A) is related to the spoken word “dog” (B; relation AB), and the picture of the dog (A) is related to the written word “DOG” (C; relation AC), then without any additional training, the learner will relate the written word “DOG” (C) to the spoken word “dog” (B; relation CB). These untrained relations are referred to as *derived relations.*

Stimulus equivalence is a composite phenomenon, made up of three critical components: *reflexivity, symmetry,* and *transitivity* (Sidman et al., 1982). Reflexivity refers to the stimulus-stimulus relation in which a stimulus is selected that is identical to the sample stimulus, in the absence of previous training or reinforcement (A=A; Cooper et al., 2007). For example, when presented with a picture of an apple, the learner will select an identical picture of an apple, without any training or reinforcement for this response. Symmetry is a stimulus-stimulus relation in which the learner exhibits a response that demonstrates reversibility of the sample and comparison stimuli in the absence of previous training or reinforcement (A=B, then B=A; Cooper et al., 2007). For example, if a child is taught to match a picture of an apple to the written word “APPLE,” then the reverse relation is also learned (i.e., the child matches the written word “APPLE” to the picture without any reinforcement or training). Transitivity is the result of training two (or more) stimulus-stimulus relations each containing partially overlapping stimuli, with the result of a new stimulus-stimulus relations emerging between the nonoverlapping stimuli in the absence of additional training/reinforcement (A=B, and B=C, then A=C; Cooper et al., 2007). For example, if a child is taught to match a picture of an apple to the written word “APPLE,” and to select the written word “APPLE” after hearing the vocal stimulus “apple,” then the untrained relation of selecting the *picture* of an apple after hearing the vocal stimulus “apple” will emerge without direct teaching.

EBI in the research literature. Rehfeldt (2011) conducted a literature review of all articles that included “derived stimulus relations” or “stimulus equivalence” in the *Journal of Applied Behavior Analysis* (JABA) from 1992–2009, resulting in a total of 26 articles. She found that 46% of studies included participants diagnosed with developmental disorders (e.g., Down syndrome, intellectual disabilities, brain injury, ASD); 19% included typically developing children with academic difficulties; 31% included individuals without educational deficits or clinical concerns, and 4% included participants with a clinical disorder (i.e., pathological gambling). The majority of studies examined the acquisition of basic vocabulary and reading skills. Fifty-four percent of the studies used pictorial/textual stimuli; 19% used numerical stimuli; 12% used randomly configured stimuli, and 15% used monetary-related stimuli. In terms of sense modality, 62% of the studies used all visual stimuli and 38% included both auditory and visual stimuli. None used stimuli pertaining to other sensory modalities (Rehfeldt, 2011). Additionally, 65% of studies used an automated method of training, while 35% used a tabletop method. All studies in the review used MTS to evaluate the development of untrained relations. Sixty-five percent of the studies assessed some form of generalization and 35% demonstrated that emergent stimulus relations generalized to stimuli that were related to the original training stimuli. Only three studies (12 %) assessed maintenance of the emergent skills. Results illustrate the potential of EBI to teach numerous skills to different populations, as well as important considerations when constructing an EBI protocol.

Variations of EBI. Several studies have explored differing trial arrangement procedures for implementing EBI and have found that some arrangements produce superior results in terms of number of relations mastered and efficiency of instruction.

Match-to-sample. The most common assay in EBI is "match-to-sample" (MTS). In essence, a "sample" stimulus is presented, and the participant is required to emit a response that demonstrates attention to the stimulus (e.g., touching the sample stimulus). Next, the "comparison" stimuli are presented, which consist of one stimulus that is programmed for reinforcement as well as other "distractor" stimuli. For example, a written Greek letter name (sample stimulus) appears in the middle of a test screen and then Greek letters (the comparisons) surround the sample (Sidman & Tailby, 1982). Reinforcement will only be delivered if the learner selects the Greek letter corresponding to the sample written word. Despite the popularity of MTS, other arrangements have been shown to be effective as well.

Four-stage equivalence model. Sidman et al. (1982) conducted a study with typically developing children to evaluate the effects of a conditional discrimination procedure to produce stimulus equivalence.

Conditional discrimination procedures establish a conditional relation between stimuli, forming an "if…then" rule (e.g., if A1, then B1 or if A2, then B2). For example, in a teaching arrangement, written Greek letter names appear in the middle of the test screen surrounded by comparison stimuli of Greek letters. If sample A1 "delta" is presented, then a response toward δ (comparison B1), but not χ (comparison B2) or ε (comparison B3) will be reinforced. If sample A2 is "gamma," χ will be reinforced if selected, but not δ or ε, if sample A3 "epsilon" is presented, only a response toward ε will be reinforced. (This stimulus display is called "simultaneous matching" because the sample stimuli remain present throughout the entire trial.) If MTS is also generated by this kind of training, then the stimuli become equivalent members of a stimulus class rather than related only by conditionality. In other words, A1 and B1 become members in one equivalence class, A2 and B2 in another, and A3 and B3 in a third.

In Sidman et al. (1982), the participants either had to select the corresponding stimuli or vocally answer with the correct response. First, participants were provided with a pre-training to acquaint them with the procedures using familiar stimuli (hue samples and comparisons). Secondly, pre-tests assessed identity matching of all Greek letters to be used. Third, participants were taught three sets of relations AB, AC, and DC. Fourth, performance was evaluated on the six sets of relations (DB, BD, AD, BC, CB, and CD) in addition to participants' oral responses to B, C, and D stimuli. Participants were required to demonstrate mastery at each step before they could move on. Following teaching, tests were conducted in the form of MTS and oral naming. A four-stage equivalence model was implemented, in which unreinforced probes (of relations DB/BD) were inserted within intermittently reinforced baseline trials (of relations AB, AC, and DC). This assay is called the four-stage equivalence model because conditional relations within the four sets of stimuli (A, B, C, and D) were required for the emergence of DB and BD to occur. Additionally, three-stage equivalence probes were implemented, which required subjects to learn conditional relations within three sets of stimuli (A, B, and C) for the emergence of BC and CB. Symmetry probes then tested performance on the DC relation, followed by oral naming tests for B, C, and D.

Sidman et al. (1982) found that six of the eight participants responded correctly on the six novel sets of conditional discriminations that were not directly taught. This study illustrated the efficiency of the four-stage equivalence model as after teaching the participants nine sample-comparison relations, 18 new stimulus relations and nine oral naming relations successfully emerged. In addition, the efficiency of the teaching model increased as the classes expanded. In other words, when three relations were taught, six additional relations were derived without direct instruction. Then with the addition of a single member to each stimulus class, fifteen novel relations were derived. The results of this study indicated that the "if…then" relations functioned to produce equivalence relations similar to an MTS assay and also demonstrated the applicability and efficiency of the four-stage equivalence model to train novel performance. Practitioners interested in teaching generative behaviors would benefit from incorporating this model into their instructional practices.

Simultaneous and simple-to-complex training. Another study that evaluated variations of EBI was conducted by Fienup, Wright, and Fields (2015). Their research compared the effects of two training protocols on creating academically applicable equivalence classes among 43 undergraduate psychology students. Both simultaneous (SIM) and simple-to-complex (STC) arrangements have been shown to successfully generate equivalence classes; however there had never been a direct comparison between the two protocols before this study. In the SIM condition, all the relations that were programmed for direct teaching were taught, then tests for derived relations were conducted. In the STC condition, training and derived relations testing trials were interspersed after each prerequisite skill was mastered (e.g., A/B taught, then B/A tested, followed by B/C taught and relations C/B and A/C tested).

Experiment 1 investigated the development of three-member equivalence classes comprising the name, illustration, and function of neuroanatomical structures (e.g., amygdala). The 3-STC group averaged 27.5% correct responses and 3-SIM group averaged 26.9% correct, with no significant difference between the groups. Five participants in the 3-SIM group exhibited a delayed development of the equivalence classes, requiring remedial training to reach mastery. In contrast, all members of the 3-STC group demonstrated immediate emergence of the untrained relations. Despite these results, efficiency measures (i.e., minutes required to form equivalence classes, minutes to demonstrate immediate emergence for three-member equivalence classes across protocols) revealed little difference between the conditions (Fienup et al., 2015).

Experiment 2 examined the development of four-member equivalence classes. The 4-STC group exhibited an average of 26.6% correct responses and 4-SIM group displayed an average of 28% correct responses. Additionally, a similar result in experiment 2 was demonstrated with immediate versus delayed emergence. While all participants in the 4-STC protocol exhibited immediate emergence, only 5 of the 12 in the 4-SIM did. Remedial training was implemented for the remaining 7 participants who failed the derived relations probes and post-tests. Overall, it is important to note that the difference in training protocol and class size both affected the immediate emergence of equivalence classes (Fienup et al., 2015).

One-to-many and many-to-one. Arntzen et al. (2010) taught music skills to a 16-year old male with ASD using an MTS teaching procedure. He was taught four 4-member classes, comparing two different structures of one-to-many (OTM; AB and AC) and many-to-one (MTO; AB and CB). This study sought to compare the two methods through a conditional discrimination procedure to teach music skills. Stimuli included major/minor chords written in Norwegian (A), written in Vietnamese (D), dots for chords on the piano keys (B), and chords written as notes (C). One of each of the chords was taught either using the OTM or MTO method (Arntzen et al., 2010).

A pre-test was conducted to assess baseline performance of untrained relations. Following training, emergent relations were tested and then additional training was conducted to expand classes further. Retraining was implemented if mastery for equivalence was not met. This consisted of mixing all the relations with feedback, slowly reducing the feedback, and administering another test phase. For the MTO structure, two sets of stimuli had to be retrained, whereas the OTM structure only required one retraining. These findings suggest OTM may be more effective in producing emergent relations. Results also indicate the importance of including possible retesting and retraining to determine at the point in which equivalence will emerge. This study illustrated the efficacy of an MTS procedure with individuals with ASD for teaching music skills (Arntzen et al., 2010).

In summary, EBI has been shown to successfully facilitate untrained behavior for several different kinds of skills, including music (Arntzen et al., 2010; Hayes, Thompson, and Hayes, 1989; Perez & de Rose, 2010), thus increasing the efficiency of teaching and allowing for a greater depth of the material to be taught. In addition, prior studies have found that the instructional arrangement can impact the formation of untrained relations. For example, the four-stage equivalence model, the STC protocol, and the one-to-many format appear to be superior to other variations of EBI.

Using Equivalence-Based Instruction to Teach Music Skills

Although EBI has been used to teach a variety of skills to children with and without disabilities, there have been only a few studies addressing its use to teach music skills. Before discussing this literature, it should be noted that the research in both stimulus equivalence and music instruction conventionally use capital letters to denote class names and musical notes, respectively. To assist the reader in discriminating which is being referred to throughout this document, musical notes will be encased in parentheses and class designations will not.

Hayes et al. (1989) evaluated whether a compound stimulus consisting of elements from two separate equivalence classes could result in novel patterns of piano playing. Nine undergraduates with no previous musical training were randomly assigned to one of three groups: timing-alone, placement-alone or timing-plus-placement. Timing-alone included playing in the correct rhythm designated by the metronome; placement-alone included finger placement on the correct key. Timing plus placement consisted of both performances. In experiment 1, pre-tests of playing the keyboard and the equivalence relations were assessed. Conditional discrimination training and equivalence testing was then implemented across a varying number of sessions, depending on the experimental condition and the number of trials needed for each participant to acquire the discriminations. Correct responses were followed by a green light and incorrect responses were followed by a red light. All relations were assessed over two blocks of 10 trials each.

Six equivalence classes related to timing were generated for the timing plus placement, and timing-alone groups. Six different rhythm patterns were used, all played using the same pitch and tempo. Class A stimuli were the pattern of auditory stimuli equaling four beats. The B stimuli were the timing patterns from class A as musical notes using a combination of quarter notes, half notes, and whole notes. The C stimuli consisted of the written terms which were used to represent each of the notes (quarter note, half note, whole note). Twelve discriminations were trained, which included six AB and six AC relations. The D stimuli were musical staffs with a mark, which paralleled notes above "middle-C" consisting of “F,” “G,” “A,” or “B”. The E stimuli consisted of four white keys with two black keys on a keyboard related to the marks on the staff from “D”. The F stimuli referred to the four fingers on the right hand (index, middle, ring, and thumb) which signified the finger to play the consequent keys with. The G stimuli represented the letter names of the notes on the staff in D (Hayes et al., 1989).

To evaluate participants’ baseline ability to play keyboard, a pre-test was conducted in which a modified keyboard, metronome, and sheet music were presented along with the instruction to play. Five different scores of music were given, which included 12 pitches with the four notes from the D stimuli, also combined as one of the three note types (quarter, half, and whole). Following training of timing and placement, a post-test was conducted. Correct responses required playing the note corresponding to the antecedent stimulus presented on that trial and holding it down for the appropriate number of metronome beats (Hayes et al., 1989).

Several interesting findings were noted. First, the number of trials to reach mastery criterion decreased across subsequent training sequences for all subjects. For example, subject 1 required five trials to master the first score, but only required two trials to master the final score, demonstrating an acceleration in learning rate as a result of exposure to the procedure. Second, participants who were exposed to the timing-alone condition could time their playing correctly, and participants in the placement-alone condition could only play the note placement correctly. Only participants in the timing-plus-placement group engaged in novel patterns of activity during the posttest when shown sheet music that included both timing and placement combinations. These findings support the hypothesis that keyboard performance was based on the learned equivalence relations. However, the experimenters questioned whether the obtained results were indeed due to the teaching arrangement or if they could be explained by practice effects, as the participants had shown improvement across different sheet music without explicit feedback. In other words, the subjects may have needed practice to achieve fluency. Additionally, verbal descriptions of relations appeared to enter a class more easily than nonverbal. As participants were verbal adults, Hayes et al. (1989) points out they may have been reading the note and timing names while pointing to them during training. This ease may refer to more formally verbal stimuli leading to enhanced playing of the keyboard. To address this question, they went on to experiment 2 (Hayes et al., 1989).

Experiment 2 was conducted to evaluate whether playing would occur despite the removal of experimenter-provided names from the equivalence classes. Nine new undergraduates were randomly assigned to one of three groups: no timing names, no placement names, and no names. All three groups were taught both timing and placement classes from experiment 1. No timing names referred to the same timing and placement classes as in experiment 1, but the class C relation was not used in training or testing (i.e., quarter, half, and whole). No placement names included the same timing and placement classes from experiment 1, but the class G relation was not used in training or testing (i.e., F, G, A, and B). No names had neither class C relation nor class G relation involved in training or testing. Procedures were identical for pre- and post-tests, as in experiment 1 (Hayes et al., 1989).

All subjects played the keyboard despite removal of names, whether one or both sets were removed. However, more trials were required to reach criterion for those in the no names group than in no timing names and timing and placement from experiment 1. It is still unclear whether experimenter-provided names aided in keyboard playing. It is possible, though, that the addition of names did allow for faster acquisition of relations and keyboard playing. Overall, these experiments demonstrated that novel musical performances can be taught through novel combinations of equivalence class members (Hayes et al., 1989).

Previous EBI studies to teach piano skills have all differed in teaching arrangements and stimuli. Moreover, they did not evaluate if the same behavioral function transferred to separate members of established equivalence classes. Griffith, Ramos, Hill, & Miguel (2018) expanded the literature on EBI for piano skills by teaching six undergraduate women to play and identify music notation of musical chords. In Experiment 1, researchers implemented an auditory-visual MTS procedure to train three chords ("C-major," "G-major," or "F-major"), in addition to three categories of visual stimuli. Visual stimuli included a textual representation of the chord name (B), a picture of the keyboard with red dots indicating the correct keys to play (C), and the music notation (D). Participants were directly taught the correct finger placement on the keyboard, then given tests for the emergence of 11 untrained relations (labeling music notation [DE], matching textual note and piano visual [BC], and playing the chord on the piano when given textual representation or musical notation [BF, DF]). Participants were then exposed to sequenced generalization tests, which included playing a song on the piano (e.g., *Amazing Grace*) when given chords either as textual representation (BF) or musical notation (CF). In Experiment 2, the picture representation of the keyboard (C) was removed from teaching, which resulted in the training requiring half the time. Results indicated that EBI was effective in producing novel piano skills and teaching adults to play a song.

Building upon the procedures implemented by Griffith et al., (2018), Hill (2015) investigated the use of EBI to teach piano skills and assess novel piano playing with four typically developing children and two children with ASD. An OTM teaching structure was used for all MTS tasks. The songs *Mary Had a Little Lamb* and *Hot Cross Buns* were used to assess novel piano playing visually represented as either letters or musical notation. All six participants reached mastery levels for novel piano playing for two different songs, though three of the six participants required remedial training. Hill (2015) also conducted a melodic probe following equivalence training and testing. The auditory stimulus of playing the note was presented and participants were asked to label the note, select the corresponding note from an array, and reproduce the response of playing the same note. Both participants with ASD exhibited above 89% accuracy for the melodic probes, whereas the typically developing children scored variably, with three of the four only scoring above 89% for the reproduction of playing the correct note and but not the other tested relations. The results with the children with ASD have not been able to be replicated with other participants (Hill, 2015).

Summary

Previous studies have illustrated the efficacy of EBI in teaching new skills to several populations (Rehfeldt, 2011). The four-stage equivalence model (Sidman et al., 1982) and STC training protocol (Fienup et al., 2015) were both shown as efficient teaching methods in producing emergent relations. Additionally, several studies have demonstrated the success of EBI in teaching music skills (Hayes et al., 1989; Arntzen et al., 2010; Perez and deRose, 2010; Griffith et al., 2018). The results of Hill (2015) further illustrate the potential of EBI procedures in teaching children to play and read music. However, the possible mechanisms controlling responding on melodic probes warrant further study.

A possible explanation that could have led to the melodic probe results in Hill (2015) was the selection of the notes “C,” “D,” and “E,” which are less easily differentiated on the staff and piano keys (both visually and auditorily) due to their similar frequencies. Other research has found that the discriminability of stimuli can impact responding. For example, in a parametric analysis of stimulus presentation, Catania (2013) found that pigeons pecked at higher rates in the presence of frequencies similar to the one that was originally trained and at systematically lower rates as the frequencies gradually differed from the original. These data formed a bell curve, creating a "generalization gradient". That is, generalization was most likely to occur in the presence of similar sounds. Said another way, discrimination between sounds was less likely to occur if the sounds were similar. This concept could contribute to a better understanding of the ability to differentiate notes/frequencies when played on the piano. Previously reinforced responses on the learned keys should lead to better differentiation of the notes. In Hill (2015), notes of "middle-C," "D," and "E" are frequencies 261.63 Hz, 293.665 Hz, and 329.628 Hz respectively. These are relatively close pitches. It may be possible that the more differentiated notes "C," "E," and "G" could be better learned "by ear," as their frequencies are further apart at 261.6 Hz, 329.628 Hz, and 391.995 Hz respectively (Suits, n.d.). The purpose of this study was to replicate and extend Hill (2015) by using more differentiated notes (i.e., "C," "E," "G"), probing for response generalization (notes "D" and "F"), and assessing maintenance of skills among children with ASD.

Method

Participants

Participants included three children diagnosed with ASD. Participants were recruited from a clinic providing intensive behavioral intervention located in the southeastern United States. To be included in the study, participants were required to have had a history of learning conditional relations with both visual and auditory stimuli. In addition, they must have been able to tact letters, match novel identical stimuli, and attend to an instructor while sitting for at least 15 minutes. In addition, participants were only included if they had no prior musical training or experience playing the piano.

Bonnie was a 6-year-old girl diagnosed with ASD and hypotonia. There is no assessment data on the severity of autism symptomology. She attended first grade, completed academics at grade-level, and spoke in full sentences; she also attended dance and swimming classes in the local community. Bonnie was non-compliant when asked to vocally communicate with adults and peers and spoke in a volume below typical conversation level. She received 20 hours per week of intensive behavioral intervention. Her scores on the Assessment of Basic Language and Learning Skills (ABLLS; Partington & Sundberg, 1998) were: cooperation and reinforcer effectiveness (84.2% of skills), visual performance (100% of skills), receptive language (84.4% of skills), intraverbals (100% of skills), motor imitation (88.8% of skills), vocal imitation (85% of skills), labeling (100% of skills), syntax and grammar (100% of skills), generalized responding (50% of skills), writing skills (100% of skills), spelling (100% of skills). following classroom routines (100% of skills), gross motor skills (96.6%), and fine motor skills (100%). Bonnie experienced deficits in the areas of play and leisure (73.3% of skills), spontaneous vocalizations (77.8% of skills), social interactions (52.9% of skills), requests (68.9% of skills), reading skills (70.5% of skills), and math skills (68.9% of skills).

Todd was a 4-year-old boy diagnosed with ASD with language impairment. Following the administration of the Autism Diagnostic Observation Schedule (ADOS), Todd's overall score fell within the high range of symptom severity. He was receiving 30 hours of intensive intervention per week and did not attend school. Todd spoke in full sentences and fluently tacted letter names and letter sounds. He did not participate in any additional leisure activities. Todd scored high on the ABLLS in the areas of: visual performance (96.3% of skills), receptive language (100% of skills), motor imitation (100% of skills), vocal imitation (100% of skills), requests (93.1% of skills), labeling (87.3% of skills), spontaneous vocalizations (100% of skills), syntax and grammar (95% of skills), play and leisure (86.7% of skills), generalized responding (100% of skills), gross motor skills (100%), and fine motor skills (100%). Todd experienced deficits in the areas of cooperation and reinforcer effectiveness (78.9% of skills), intraverbals (69.3% of skills), social interactions (76.5% of skills), reading skills (41.2% of skills), math skills (65.5% of skills), following classroom routines (20% of skills), writing skills (40% of skills), and spelling (28.6% of skills).

Jasper was a 6-year-old boy diagnosed with ASD. Following administration of the ADOS, Jasper also scored within the high range of symptom severity, with severe receptive and expressive language delay. He attended first grade in a varying exceptionalities classroom, could speak in full sentences, and read at a first-grade level. However, he could not recall previous events or discriminate between "wh" questions. Jasper also required multiple prompts to stay on task and follow routine instructions.

Jasper did not participate in any additional leisure activities. He received 20 hours of intensive behavioral intervention per week to address these deficits. Jasper scored high on the ABLLS in the areas of visual performance (96.5% of skills), motor imitation (95.4% of skills), vocal imitation (100% of skills), requests (100% of skills), spontaneous vocalizations (94.4% of skills), generalized responding (91.7% of skills), gross motor skills (93.3% of skills), fine motor skills (96.4% of skills), and spelling (92.9% of skills). Jasper experienced deficits in the areas of cooperation and reinforcer effectiveness (64.5%), labeling (75% of skills), syntax and grammar (47.5% of skills), play and leisure (73.3% of skills), intraverbals (71.9% of skills), social interactions (22.1% of skills), reading skills (76.5% of skills), math skills (31% of skills), following classroom routines (55% of skills), and writing skills (75% of skills).

It should be noted that one participant (age 4) was excused from the study during the data collection process due to deficits in prerequisite skills, including attending and difficulty discerning between more complex stimuli (music notation rather than images). This participant also exhibited noncompliance behaviors (noncompliance to experimenter's instructions and not sitting in the chair appropriately). This participant passed the pre-assessments and had a history of conditional discrimination but had trouble when tested to differentiate between the three music notation stimuli. It is possible that his young age and/or more limited verbal skills may have contributed to his difficulty with the task. Further modifications may be required to be effective with younger learners such as within-stimulus prompts and more dense reinforcement schedules.

Setting and Materials

Sessions were conducted in a treatment room at an autism treatment clinic located in the Southeastern United States. Materials included a table, two chairs, stimulus cards, and a Yamaha© 76-key portable grand piano. Other materials included a video camera for recording sessions, data sheets (see Appendix A), edibles to be used as reinforcers, a treasure box with prizes, and a three-ring binder containing sheet music. Stimulus cards were images of the notes "C," "D," "E," "F," and "G" presented as either the letter or musical notation. See Figure 1 for representation of the stimuli used during training.

Experimental Design

A nonconcurrent multiple-probe across participants design was conducted to determine the effects of EBI training on novel piano playing performance and control for repeated exposure to the stimuli during the pre-test and training periods (Horner & Baer, 1978). All relations were tested before training to demonstrate that they were unknown and after training to assess the effects on acquisition, generalization, and maintenance for all participants.

Measures

Dependent variable. The dependent variable was the percentage of correct independent responses emitted by the participant during each test of the teaching session. A correct response was recorded when, in the presence of a sample stimulus, the participant selected the corresponding comparison by placing a finger a stimulus card (AB, AC, BC), verbally stating the correct answer (BD, CD), or playing the correct key on the keyboard (AE, BE, CE). Incorrect responses included selecting a comparison that did not correspond with the sample stimulus, stating an incorrect answer, saying “I don’t know,” and not responding within 10 seconds of the presentation of the sample stimulus.

Interobserver agreement and treatment integrity. A second trained observer collected data independent of the primary observer during 33.6% of total trials to assess for interobserver agreement across both trained and derived relations. An agreement was recorded if the primary and secondary observers both scored a correct, prompted, or incorrect response for the same trial. Point-by-point agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and then multiplying by 100 (Kazdin, 2011). Interobserver agreement data were collected for Bonnie during 32.9% of trials and calculated at 97.7% (range 77.8–100%). Interobserver agreement data were collected for Todd during 22.1% of trials and calculated at 98.7% (range 77.8–100%). Interobserver agreement data were collected for Jasper in 43.9% of trials and calculated at 98.8% (range 55.6%–100%). Interobserver agreement across participants averaged 98.5% (range 97.7–98.8%).

Treatment integrity was collected on 33.6% of sessions, across all conditions, to evaluate the extent to which the procedures were implemented with fidelity. A trial was scored as correctly implemented if the experimenter delivered the correct instruction, provided the correct prompt, and delivered the appropriate consequence. The score was calculated by dividing the number of correctly implemented trials by the total number of trials and multiplying by 100 (Kazdin, 2011). Treatment integrity scores across participants ranged between 98.6 and 100%. Treatment integrity data were collected for Bonnie in 32.9% of trials, resulting in a score of 100%. Treatment integrity data were collected for Todd in 22.1% of trials, resulting in a score of 100%. Treatment integrity data were collected for Jasper in 43.9% of trials, resulting in a score of 98.6%.

Procedures

Participants stayed after their regularly scheduled behavioral intervention sessions or came to the treatment center on the weekend for music lessons with the researcher. Each of these music lessons lasted one to two hours. Sessions consisted of 9 trials and were run continuously with a break every 15 minutes. A one-to-many (OTM) teaching structure was employed as previous research has demonstrated to be effective (Arntzen et al., 2010). At the end of each music lesson, participants could select a toy from a treasure box. See Figure 2 for a flow chart depicting the order in which the procedures were implemented.

Preference assessment. A multiple stimulus without replacement (MSWO) preference assessment was conducted prior to each session to establish an item that would be used as a reinforcer throughout session

(DeLeon & Iwata, 1996). An MSWO is a quick method to determine a hierarchy of preferred items by presenting potential reinforcers in an array and asking the participant to select one. After the selection, the remaining edibles are re-presented to allow a new selection. These steps were repeated until all edibles were chosen and a preference hierarchy was formed. At the beginning of each trial block, researchers provided a choice between the top two preferred edibles identified during the MSWO to help prevent any possible satiation.

Pre-training. Pre-training was conducted to familiarize the participant with the procedures used during teaching and testing. One 9-trial block was conducted using common stimuli (e.g. apple, dog, balloon) to provide exposure to each type of training condition that the participant would encounter (i.e., auditory discrimination, tact, and listener trials). No feedback was provided during pre-training. Participants were required to score at least 89% correct across all three blocks of pre-training to continue in the study.

Test conditions. A number of test conditions, presented before, during and after the training, were conducted to evaluate the participants' performance on directly trained as well as derived relations. During all tests, instructions were given at the onset of the condition and no feedback was provided for correct or incorrect responding.

Sequenced generalization pre/post-tests. Tests were conducted to assess each participant's ability to play sequences of notes on the piano prior to and after training letters (BE) and sheet music (CE). At no time during the study were the participants directly taught to play sequences of notes; they were only directly taught to play one note at a time. Thus, this was a test of generalization to stimuli likely to be encountered in more traditional musical training. During these tests, participants were shown a series of letters or musical notation and asked to play them on the piano (BE and CE). The pre/post-test consisted of two different note arrangements (see Appendix B). The order of presentation of note arrangements was counterbalanced across participants to control for sequencing effects (see Table 1).

Visual-visual pre/post-tests. This condition was designed to test the selection of letters (B) in presence of musical notation (C) and vice versa. For each trial, a sample stimulus (C or B) was presented and three comparison stimuli were placed below (B or C). On each trial, participants were told a letter corresponding to one of three stimuli placed in front of them. After a selection was made (letter or musical notation) that corresponded to the letter they heard, the trial ended. Participants were required to score at or above 89% correct across one 9-trial block to move onto the next phase. Remedial training was provided if the mastery criterion was not met.

Auditory-visual pre-tests. This condition was designed to assess for performance on the stimuli to be directly trained (AB and AC). No post-test was conducted for these relations as participants needed to reach mastery criterion in order to move on to the next phase. On each trial, participants were told a letter corresponding to one of three stimuli placed in front of them and instructed to select the letter or musical notation that corresponded to the letter they heard.

Textual pre-/post-test. This condition was designed to assess the correct vocal response (D) when given the musical notation (C). On each trial, participants were shown a picture card with musical notation and instructed to say the corresponding note.

Piano pre-test. This condition was designed to evaluate correct finger placement on the piano (E) after hearing a vocal instruction to play the note (A). A pre-test of the AE relation was conducted and then directly trained at a later time during the study. No post-test was given, as mastery was required to move onto the next phase. At the onset of this condition, the experimenter provided instructions to the participants. No feedback was given for correct or incorrect responding.

Transfer of function post-test. This condition was designed to evaluate the performance of correct key selection on the piano when given a stimulus that was related via equivalence (BE and CE). A sample stimulus was presented to the participant with the instruction to play the corresponding key on the piano.

Training. The experimenter used a progressive prompt delay during training (Touchette, 1971) in which a gestural prompt was used to signal the correct response. The prompt was delayed by 0 seconds, 2 seconds, and 5 seconds following the instruction over successive trials. The criterion requirement to increase the prompt delay was two consecutive trial blocks at 89% or above, which included both independent and prompted responding. Three consecutive errors within one block resulted in a return to the previous prompt level. At the end of each trial block, a vocal acknowledgement (e.g., "All done" or "You're finished") was provided without any error correction, prompts, or consequences.

Differential reinforcement was used for both independent and prompted responses (Karsten & Carr, 2009). Prompted correct responses resulted in mild praise (e.g., "Yes, that's right") and independent responses were followed with a token and enthusiastic praise (e.g., "Yay! Awesome job!"). Cells on the token board were used to designate when backup reinforcement (i.e., edibles) would be delivered. A FR1 schedule of reinforcement was used throughout the training. Following mastery, the reinforcement schedule was thinned to FR9 for one block, followed by withholding reinforcement for another block. Prompt and reinforcement fading replicated that of Hill (2015) described above.

Training 1. This condition was designed to teach the participants to respond correctly by selecting the related stimulus [letter for note (B), musical notation (C), playing the note on the piano (E)] when given an auditory sample stimulus (A). The initial training phase consisted of training AB and AC in mixed order. The experimenter first gave instructions on how to proceed. On each trial, the experimenter told the participant a letter and placed three stimuli on the table. The correct answer was prompted; prompts were faded systematically over trials. Once participants reached the mastery criterion and maintained independent responding across the two 9-trial blocks with reinforcement thinned as described above, they moved on to the next phase.

Remedial training 1. This training was implemented if the participant failed to pass the CD, BC, or CB post-tests. The remedial training consisted of a return to the training of AB and AC relations. As such, this training was identical to training 1.

***Training 2*.** This condition was designed to train participants to play corresponding notes on the piano (E) when given the auditory stimulus (A). Participants progressed to this condition only after achieving mastery for CD, CB, and BC post-tests. The experimenter said a letter ("C," "E," or "G") and modeled the finger placement. The participant had to reach mastery criteria and maintain independent responding across two 9-trial blocks with reinforcement thinned as described above to complete this condition.

***Remedial training 2*:** This training was implemented if a participant failed the transfer of function post-test. The remedial training consisted of a return to the training of the AE relation. As such, this training was identical to training 2.

Probes.

Melodic probes. This condition was designed to evaluate the participants' ability to tact the note when given an auditory stimulus of the note played on the piano (FD), select the picture of the note on a staff when given the auditory stimulus (FC), and play the correct note on the piano after the stimulus was provided (FE). For the FD probe, the experimenter played a note on the piano, and participants were to tact the note while facing away from the experimenter and keyboard. For the FC probe, participants still faced away from the keyboard and experimenter. The experimenter played a note on the piano and the participant selected from an array of three comparison stimuli. For the FE probe, the participant sat with the experimenter at the keyboard, and the participant's view of the keyboard was blocked. The experimenter played a note on the piano and the participant then responded by playing a key. The experimenter provided instructions and no feedback was given.

Generalization probes. This condition was designed to evaluate the participants' ability to derive, tact, and play a novel note when given the musical notation of the untaught note. The untaught notes were the two notes between the trained notes on the keyboard ("D" and "F"). For the untaught notes, probes consisted of the same equivalence relations tested for taught notes: BC, CD, BE, and CE. The probes consisted of 10-trial blocks and took place following post-test measures and preceding maintenance probes.

Maintenance probes. This condition was designed to evaluate retention of the trained and untrained relations during a follow-up session that occurred at least one week following the completion of all training and testing phases. This session was identical to the transfer of function post-test and sequenced generalization tests.

Results

Figure 3 depicts the percentage of correct responses for sequenced generalization tests, auditory-visual MTS, visual-visual MTS, textual tact, transfer of function, and melodic probes across participants. Scores for pre-tests ranged as follows: sequenced generalization pretests 0–22%; AC: 0–33%; AE: 0–22%; BC: 22–44%; CB: 11–55%; CD: 0–11%; BE: 0–22%; CE: 0%. All participants scored at or above mastery criteria for the AB pre-test, as knowing the alphabet was a required prerequisite skill. Participants required an average of 252 trials (180–333) during training 1 (AB/AC mixed) to achieve mastery criterion for the AC relation. All participants scored at or above mastery criteria for the CD, BC, and BC post-tests without any remedial training. For training 2, participants required an average of 267 trials (180–333) to meet mastery criteria for the AE relation. No participants required any remedial training and passed all post-tests for BE, CE, and the sequenced generalization post-tests.

Todd

Directly trained relations. Todd scored at 0% for pre-tests of relations AC and AE, and at 89% for AB. Todd met mastery on training 1 (AB/AC mixed) in 180 trials. For training 2, Todd required 333 trials to meet mastery criteria for the AE relation.

Derived relations. Todd scored 0% on the sequenced generalization pre-tests 1 and 2, in addition to relations BE and CE. Additionally, he scored 11% on the CB and CD relations, and 22% on BC. Following training 1 (AB/AC mixed), Todd scored 89% on the CD post-test, and 100% for the BC and CB post-tests. After training 2, he scored 89% for both transfer of function post-tests (BE and CE). Todd scored 89% for sequenced generalization post-test 1 (BE) and at 100% for sequenced generalization post-test 2 (CE).

Generalization. On melodic probes, Todd scored 22% on FD, 33% on FC, and 22% on FE. Todd also scored low on generalization probes on untaught notes of "D" and "F" at 40% on both BC and CD, and 0% on BE and CE.

Maintenance. Todd's maintenance probes stayed at 100% for BE and CE. Todd scored 89% and 78% on sequenced generalization test 1 (BE) and 2 (CE), respectively. However, it is important to note that Todd skipped letters which were scored as incorrect.

Bonnie

Directly trained relations. Bonnie scored variably throughout pre-tests at 22% on AE1 and AE2, 33% on AC1, and 44% on AC2. Bonnie also scored 100% on both AB pre-tests. Bonnie met mastery on training 1 for the AC relation in 243 trials. For training 2, Bonnie required 288 trials to meet mastery criteria for the AE relation.

Derived relations. Bonnie scored 0% on sequenced generalization pre-tests 1 and 2, 0% on CD and CE, 44% on BC, and 55% on CB. Following training 1, she scored 100% on all three post-tests of CD, BC, and CB. Following training 2, she scored 89% on the BE post-test and

100% on the CE post-test. Bonnie then scored 89% on both sequenced generalization post-tests.

Generalization. Bonnie scored low on melodic probes at 55% on FD and 22% on FC; however, she did score higher at 78% on the FE probe. For generalization probes on untaught notes of “D” and “F,” Bonnie scored 60% on BE, 80% on CE, 90% on BC, and 100% on CD.

Maintenance. Bonnie maintained skills, scoring 100% on BE and CE post-tests, in addition to both sequenced generalization post-tests.

Jasper

Directly trained relations. Jasper responded variably throughout pre-tests at 0% on AE2, 11% on AE1, and 33% on all AC pre-tests. Jasper also scored 100% on all AB pre-tests. Jasper met mastery on training 1 for the AC relation in 333 trials. Jasper required 180 trials in training 2 to meet mastery criteria for the AE relation.

Derived relations. Jasper scored 0% on sequenced generalization pre-test 1 (BE) and 22% on pre-test 2 (CE). Jasper’s pre-test performance was variable at 0% on CD and CE, 33% on CB, and 44% on BC. Following training 1, Jasper scored 100% on all three post-tests of CD, BC, and CB.

After training 2, he scored 100% on all transfer of function post-tests (BE and CE) and sequenced generalization post-tests 1 and 2.

Generalization. For melodic probes, Jasper scored low at 44% on FD, and 55% on FC and FE. Jasper scored low on generalization probes as well, at 50% on BC, 0% on CB, and 10% on BE and CE.

Maintenance. Jasper scored high at 100% for both sequenced generalization tests and at 78% for maintenance probes of BE and CE. It is possible that noncompliance behaviors may have artificially deflated his performance for initial probes of BE and CE.

Discussion

The purpose of this study was to systematically replicate and extend Hill (2015) by evaluating the use of a stimulus arrangement procedure to facilitate equivalence relations between auditory stimuli (A), textual notes (B), and musical notation (C) for musical notes of "C," "E," and "G". Using these procedures, children with ASD were directly taught three relations: auditory-textual (AB), auditory-music notation (AC), and auditory-playing (on piano; AE). Emergence of novel piano skills were tested, and participants demonstrated mastery for five untrained relations: textual-music notation (BC), music notation-textual (CB), music notation-vocal response (CD), textual-playing (BE) and music notation-playing (CE). All three participants passed post-tests and transfer of function tests without any remedial training.

The results of this study are inconsistent with Hill (2015), in which three participants in the previous study required remedial training before passing post-tests. However, Todd, Bonnie, and Jasper required more trials to reach mastery criteria than participants in the previous study. It is possible the increased number of required training trials were due to the younger age and less sophisticated verbal repertoires of the participants in this study. Additionally, the prompt delay procedure may have contributed to the high number of trials to criterion for all participants. Jasper would attempt to touch the card/key before the therapist could prompt (e.g., "I want to touch it first!"). This could have artificially increased the trials to criterion. A different type of prompting may have been warranted due to noncompliance behaviors (e.g., within-stimulus prompts). Of note, Todd and Bonnie played the sequence and intervals between keys correctly (i.e., each note separated by one white key) during training 2 (AE), but their finger placement on the keyboard was incorrect. These errors resulted in a return to the previous prompt level, thereby increasing the overall number of trials to mastery criterion. Hill reported similar results with two participants. Both erred during initial post-tests due to initial incorrect finger placement, although the sequence was correct.

These observations suggest Todd and Bonnie were responding to the order of the notes instead of the location along the keyboard. This error pattern is unfavorable as the location of “C,” “E,” and “G” are white keys associated with the pairs of black keys along the keyboard (i.e., “C” and “E” surround the pair of black keys). Future researchers may want to examine blocking off one portion of the keyboard during training, reducing the opportunity to make errors, then gradually increasing the number of keys shown as training progresses.

Playing Sequences of Notes

Results of the sequenced generalization probes, in which participants were tested on the ability to play a series of notes without ever having been taught to do so, were consistent with the results obtained by Hill (2015). All participants in both studies scored at or above mastery levels, where letters (B) or musical notes (C) were provided in novel order. While the arrangement of notes in Hill’s study were comprised of the recognizable songs *Mary Had a Little Lamb* and *Hot Cross Buns,* the musical notes in this study were “C,” “E,” and “G,” placed in novel order. This is notable because familiarity with a melody prior to instruction can increase accuracy in piano playing performance (Frewen, 2010). Frewen (2010) found that after a brief training, children familiar with a melody played more accurately than children unfamiliar with the melody. It is possible that familiarity can assist in detecting errors (Frewen, 2010). In addition to the recognizable melodies in Hill’s study, the combination of “C,” “D,” and “E” notes appear in many other children’s songs (e.g., *Itsy Bitsy Spider*). Because the participants in the current study responded accurately to notes that do not commonly occur together in children’s songs and the resulting melodies were unfamiliar, this lends further credibility to the efficacy of the EBI procedures because it rules out familiarity as an explanation for the obtained results. Nevertheless, researchers may find it beneficial to further investigate the effects of using familiar melodies combined with EBI.

Melodic Probes

No participants performed at mastery for the melodic probes, i.e., auditory note-music notation (FC), auditory note-vocal response (FD), auditory note-playing note (FE). However, all participants demonstrated transfer of function between the auditory (A), textual (B), and musical notation (C) for each musical note within each class. In other words, the presentation of B and C resulted in reading music notation and the presentation of A, B, and C all resulted in playing the correct key on the piano.

Generalization probes for Todd and Jasper were low ranging between 0-50% for all four probes of (BC, CD, BE, CE). Interestingly, though Bonnie scored at 60% for playing when given the textual representation, she scored much higher at 80% for playing when given the musical notation (CE), 90% for matching the textual to the music notation (BC) and 100% for reading the musical notation (CD). Maintenance probes were high across participants at 100% for both BE and CE for Todd and Bonnie, with Jasper scoring at 78% for both. However, both Jasper and Bonnie scored at 100% for sequenced generalization tests 1 and 2. Todd scored 78% for sequenced generalization test 1 (BE), but at 89% for sequenced generalization test 2 (CE). Following the maintenance probe of sequence generalization test 1, Todd asked for the paper to be brought closer to his eyes, vocalizing “I can’t see it”. Todd performed with higher percentage correct for sequenced generalization test 2 when the paper was brought closer to his eyes. Todd's possible vision problems could have influenced his lower performance for the initial sequenced generalization test and improved performance for sequenced generalization test 2.

Results for the melodic probes (FC, FD, and FE) do not correspond with previous research. In Hill (2015), two participants scored at or above mastery criteria (89%) on the melodic probes. This suggests that the use of more differentiated notes "C," "E," and "G" were not more successfully discerned for auditory melodic probes than the closer notes used in previous studies. Interestingly, though Bonnie scored lower for the FC and FD probes (33% and 55% respectively), she scored at 78% for the FE probe of imitating note playing following only the auditory stimulus of the note being played. Though she missed the first two, she scored correctly on the following probes suggesting that she may have self-corrected on subsequent trials. A similar result was observed in Hill with three of the four typically developing participants scoring at or above 89% for the FE probe only.

The purpose of the melodic probe was to investigate the possibility that exposure to musical stimuli leads to learning to identify melodic sounds without additional training. Additionally, the intention was to investigate whether the correct note playing was potentially under the control of the visual stimulus (sight of the finger placement on the key) or the auditory stimulus. Hill (2015) determined that the responding of the four typically developing children was likely under the control of the visual stimuli as they could not reliably tact, match to musical notation, or imitate note playing when they could not see which note was played. Furthermore, she concluded that the responding of the participants with ASD was possibly under the control of the auditory stimuli, as both scored at or above mastery for all melodic probes. The results of the current study do not corroborate these findings. However, few studies have included a melodic testing component and future research should investigate the mechanisms possibly controlling musical responding to further develop and refine teaching procedures.

Bonnie's performance on the novel notes "D" and "F" is notable. This illustrates the successful implementation of the *setup* and *difference principle*. The setup principle states that when teaching a concept, exemplars should share the most possible irrelevant features (Engelmann & Carnine, 1982); in this way, learners can better distinguish what it is they are to attend to. For example, for different musical notes stimuli should be identical except for the concept that is being taught, such as the location of the note on a line or space. All other aspects would be identical (e.g., note type, clef, and staff). A non-example of this would be using two different note values (e.g., quarter note and half note). This would not signify the setup principle, as the learner could attend to the many irrelevant features which would distract from the concept. The difference principle is similar in that it shows the limits and boundaries of a concept, including similar examples and non-examples, which only differ in the critical feature (Engelmann & Carnine, 1982). This is also most effective when the stimuli are presented beside each other for better comparison (Watkins & Slocum, 2003). By following these principles, one should expect to derive untaught notes of "D" and "F" both on the staff and on the piano, as they appear sequentially in order on both stimuli and alphabetically. All aspects of the notes were the same in teaching the location of the note on the staff, including how the note appeared on the staff. The only difference was the location on either a line or a space. If this is all true, participants should be able to derive the untaught notes. Bonnie's ability to label musical notation, match the textual to musical notation, and play correct notes on the piano when presented with untaught stimuli further illustrates the potential and success of EBI teaching procedures.

Transfer of Function

Consistent with Hill (2015), all participants exhibited transfer of function. Transfer of function occurs when other members of the same class acquire the function of a given stimulus without any direct teaching (Hayes & Hayes, 1992; Perkins, Dougher, & Greenway, 2007). This phenomenon has been demonstrated in previous studies across differing populations and with various behaviors such as clapping and waving (Barnes, Browne, Smeets, & Roche, 1995), following picture schedules (Miguel, Yang, Finn, & Ahearn, 2009), and playing musical chords on a piano (Griffith et al., 2018). Transfer was observed in the current study as piano playing (E) occurred when letters (B) and music notation (C) were presented without any additional teaching. This demonstration of transfer of function to produce novel piano playing in young children strengthens the external validity of EBI procedures.

Contribution of Verbal Behavior

Verbal mediation may also be a key component in investigating the mechanisms controlling participant responding (Santos, Ma, & Miguel, 2015). Consistent with results from Hill (2015), all three participants in the current study engaged in vocalizations (e.g., tacting during the visual-visual

MTS) during training and testing that corresponded with the stimuli being selected (i.e., the note/letter names of "C," "E," and "G"). Todd, Bonnie, and Jasper all read the notes aloud before playing them on the keyboard during training and post-tests. Todd and Jasper also read aloud before playing during the sequenced generalization tests and it is possible Bonnie was engaging in covert verbal mediation during this task. This suggests that the participants' verbal behavior may have played a role in their performance. Previous studies have questioned if verbal mediation, specifically *naming*, is necessary for the development of stimulus equivalence (Sidman, Willson-Morris, & Kirk, 1986).

Naming involves a bi-directional relation between a class of stimulus and the verbal behavior they occasion (Horne & Lowe, 1996). Previous research has explored the relationship between naming and the success of MTS tasks (Horne, Hughes, & Lowe, 2006; Sidman et al., 1986). An individual is said to have a generalized naming repertoire when he or she responds as both speaker (tacting) and listener (selecting the stimulus) when only one is directly trained (Santos et al., 2015). Previous research has illustrated that responding as both speaker and listener resulted in accurately sorting stimuli into classes (Kobari-Wright & Miguel, 2014;

Miguel et al., 2015; Ribeiro, Miguel, & Goyos, 2015). The opposite has also been observed, where those who could tact but not engage in listener behavior or vice versa were more inclined to fail matching tasks and/or novel categorization (Miguel, 2018). Horne et al., (2006) illustrated that the naming of arbitrary stimuli was effective in establishing stimulus classes. Participants were trained on common listener relations, and in the process of doing so, also exhibited transfer of function. The authors argued that dissimilar stimuli that evoke common speaker and listener behaviors can become equivalent. Future research should continue investigating the possible effects of naming on producing equivalence. Additionally, researchers should ensure to include sufficient measures of the verbal repertoires of participants with disabilities in relation to the stimuli used (Horne et al., 2006).

Further study is warranted on the role of verbal mediation in skill acquisition. One possible method is by examining response latency in MTS tasks. Previous research suggests that increases in response latency observed from baseline to testing are due to mediating or "problem solving behavior" (Miguel, 2018). During problem solving, individuals engage in a series of overt or covert behaviors to reach a solution (e.g., echoic rehearsal;

Miguel, 2018). It may be important to investigate factors that lead to the ability to translate this teaching format to those with more limited verbal skills, as these children would benefit from involvement in appropriate leisure activities.

Social Validity

Parents of Bonnie and Jasper have reported they will start music lessons to continue their music education, incorporating a more appropriate leisure skill into their daily lives. Both Bonnie and Jasper also expressed excitement about sessions and learning to play the piano to their peers at the center, resulting in numerous peers asking if they could also start piano. These results demonstrate the efficacy of EBI procedures; however, further investigation is required to explore the use of EBI procedures in teaching music to individuals with disabilities, especially learners with more limited skill sets.

Limitations

Though conclusions from this study are promising, procedural limitations should be considered. First, the AB portion of training 1 could be removed due to the high performance for the AB pre-test, and then be applied only to those who failed the pre-test in addition to providing only periodic probes for those who passed the pre-test, saving time during the training process. Including the AB portion also over-inflates the data for the AB/AC mixed training, as AC is the only relation to be mastered. Second, the sequenced generalization test BE (textual to playing) could be removed, as music is played using musical notation and not letters. Third, this study was implemented using stimulus cards in a table-top method rather than a computer-based program as used in Hill (2015), which potentially led to longer intertrial intervals and treatment integrity errors. Fourth, participants in this study were also younger than those in Hill and may not have had as sophisticated verbal repertoires as those in previous research, leading to slower learning and more problematic behaviors that were not observed during pre-assessments (noncompliance, attending, etc.). Fifth, Jasper was more interested in gaining access to the treasure box prize at the end of session rather than tokens or edibles selected during the MSWO. Though he continued working to earn tokens and edibles, he consistently interrupted trials and vocalized, "When do I get my prize?" It may be useful to conduct a reinforcer assessment, rather than a preference assessment, to establish reinforcers that would potentially increase motivation. Sixth, maintenance probes were collected after only 1-2 weeks following post-test mastery, and future studies should consider including additional probes that test maintenance of skills taught over a longer period.

The procedures described in the study should be replicated to further examine the use of EBI procedures for more complex music skills such as for timing training with finger placement on the piano keyboard (e.g., with a metronome), in addition to chord playing training (e.g., playing multiple notes in unison). These skills are the goal of music instruction, leading to the ability to play a whole song in time or "on beat". Skinner (1959) emphasized the importance of developing a "sense of rhythm" (p. 220). A sense of rhythm refers to proper timing of behavior. Skinner implemented the use of a teaching machine to teach children rhythmic patterns where they would tap in unison with the machine. At first, participants' responding was reinforced even if they were a little early or a little late, and then timing was slowly sharpened. Hayes et al., (1989) also investigated different methods to teach timing in addition to finger placement in piano playing among undergraduate students. It may be beneficial to examine the efficacy of different protocols for teaching appropriate timing in addition to finger placement and reading musical notation.

Next, the effectiveness of the OTM and MTO training structures should also be investigated to compare efficacy of these procedures. Horne et al. (2006) summarized the effects of the structure of the conditional discrimination training on performance. Arntzen et al., (2011) concluded that the OTM structure was more efficient is producing equivalence. However, past studies have also illustrated that children with disabilities respond more accurately when taught with an MTO training structure in comparison to OTM (Horne et al., 2006; Saunders, Drake, & Spradlin, 1999). Additionally, participants given OTM training structures were more likely to fail initial tests, requiring additional teaching sessions (Green, 1990).

Lastly, future research should explore EBI procedures in comparison to standard musical instruction such as the Suzuki method (Suzuki, 1978) and Alfred's Basic Piano method (Palmer, Manus, & Lethco, 1981). Many within the music community assert that music skills are inherent, but previous studies suggest that it can be taught through specific teaching methods (Brandt, Gebrian, & Sleve, 2012; Lehmann & Ericsson, 1997). Therefore, it is important to further investigate the mechanisms and efficacy of the teaching methods that leads to music skill acquisition. Prior research has supported EBI procedures as a powerful technology in teaching a variety of skills across varying populations. The success and potential of EBI procedures warrant further study and support to encourage implementation across settings and ensure the success of children with disabilities.

Results of the current study illustrate the success of teaching an appropriate leisure skill to children with ASD. Previous studies have illustrated various benefits of music instruction for children with ASD, such as academic improvements (Brandt et al., 2012), increased social behaviors (Eren, 2015), increased speech production (Lim et al., 2011), improved joint attention skills (Kim et al., 2008), and decreased vocal stereotypy (Lanovaz et al., 2011). Music is an activity that may be beneficial for children with ASD who frequently lack appropriate leisure skills and have more limited access to reinforcers. Previous research illustrates the importance of participation in leisure activities for children with ASD. Children with disabilities participate in fewer social activities and can become more passive in comparison to their peers (Cannella-Malone, Miller, Schaefer, Jimenez, Page, & Sabielny, 2016). Involvement in leisure activities has led to positive effects such as increasing activity level, social interactions, and community involvement (Canella-Malone et al., 2016). Despite these positive effects, engagement in leisure skills is regarded as a low priority in many schools. Leisure activities add to children's quality of life by encouraging engagement with their environments. Leisure activities are also a potential reinforcer for teaching academic, social, and functional skills (Canella-Malone et al., 2016). The benefits of involvement in leisure activity are vital for the continued development and success of children with disabilities.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, D.C: American Psychiatric Association.

Arntzen, E., Halstadtro L., Bjerke, E., & Halstadtro, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions, 25,* 129-143. https://doi.org/10/1002/bin.301

Autism Speaks. (2017). Treatments & therapies. Retrieved from https://www.autismspeaks.org/family-services/tool-kits/100-day-kit/treatments-therapies

Barnes, D., Browne, M., Smeets, P., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three to six-year-old children. *The Psychological Record, 45,* 405-430. https://doi.org/10/1007/BF03395151

Blum-Dimaya, A., Reeve, S. A., Reeve, K.F., Hoch, H. (2010). Teaching children with autism spectrum disorders to play a video game using activity schedules and game-embedded simultaneous video modeling. *Education and Treatment of Children, 33*(3), 351-370. https://doi.org/10/1353/etc.0.0103

Brandt, A., Gebrian, M., & Sleve, L.R. (2012). Music and Early Language Acquisition. *Frontiers in Psychology, 3,* 327. https://doi.org/10.3389/fpsvg.2012.00327

Canella-Malone, H.I., Miller, O., Schaefer, J.M., Jimenez, E.D., Page, E.J., Sabielny, L.M. (2016). Using video prompting to teach leisure skills to students with significant disabilities. *Exceptional Children, 82*(4), 463-478. http://doi.org/10.1177/0014402915598778

Catania, A.C. (2013). *Learning* (5th ed.). Cornwall-on-Hudson: Sloan Publishing.

Carlile, K. A., Reeve, S. A., Reeve, K. F., DeBar, R. M. (2013). Using activity schedules on the iPod touch to teach leisure skills to children with autism. *Education and Treatment of Children, 36*(2), 33-57. https://doi.org/10/1353/etc.2013.0015

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis.* Upper Saddle River, NJ: Pearson Education Inc.

Cross, I., Fitch, W. T., Aboitiz, F., Iriki, A., Jarvis, E. D., Lewis, J., Liebal, K., Merker, B., Stout, D., & Trehub, S. E. (2013). *Culture and Evolution.* In M.A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship* (pp. 542-543) Cambridge, MA: The MIT Press.

Cunningham, A. B., & Schreibman, L. (2009). Stereotypy in autism: The importance of function. *Research in autism spectrum disorders, 2*(3), 469-479. http://doi.org/10.1016/j.rasd.2007.09.006

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*(4), 519–533. http://doi.org/10.1901/jaba.1996.29-519

Engelmann, S. & Carnine, D. (1982). Theory of Instruction. New York, NY: Irvington Publishers, Inc.

Eren, B. (2015) The use of music interventions to improve social skills in adolescents with autism spectrum disorders in integrated music therapy sessions. *Procedia – Social and Behavioral Sciences, 197*, 207-213. https://doi.org/10/1016/j.sbspro.2015.07.125

Fienup, D. M., Wright, N. A., Fields, L. (2015). Optimizing equivalence-based instruction: Effects of training protocols on equivalence class formation. *Journal of Applied Behavior Analysis, 3*(48), 613-631. https://doi.org/10/1002/jaba.234

Frewen, K.G. (2010). Effects of familiarity with a melody prior to instruction on children's piano performance accuracy. *Journal of Research in Music Education, 57*(4), 320-333. http://doi.org/10.1177/0022429351178

Gaser, C., Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*(27), 9240-9245. https://doi.org/10/1523/JNEUROSCI.23-09240.2003

Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation, 95,* 260-270.

Grey, I. M., & Hastings, R. P. (2005). Evidence-based practices in intellectual disability and behaviour disorders. *Current Opinion in Psychiatry, 18*(5), 469-475. http://dx.doi.org.portal.lib.fit.edu/10.1097/01.yco.0000179482.54767.cf

Griffith, K.R., Ramos, A.L., Hill, K.E., & Miguel, C.F. (2018). Using equivalence-based instruction to teach piano skills to college students. *Journal of Applied Behavior Analysis, 51*(2), 207-219. https://doi.org/10/1002/jaba.438

Hayes, L. J., Thompson, S. & Hayes, S. C. (1989). Stimulus equivalence and rule following. *Journal of Experimental Analysis of Behavior, 52*(3), 275-291. https://doi.org/10.1901/jeab.1989.52-275

Hayes, S.C., & Hayes, L.J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist, 47*(11). 1383-1395. https://doi.org/10.1037/0003-066X.47.11.1383

Hill, K. E. (2015). Using equivalence-based instruction to teach piano skills to children (Unpublished master's thesis). California State University, Sacramento, CA.

Horne, P.J., and Lowe, C.F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*(1), 185-241. https://doi.org/10/1901/jeab.1996.65-185

Horne, P.J., Hughes, J.C., & Lowe, C.F. (2006). Naming and categorization in young children: IV: Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior, 85*(2), 247-273. http://doi.org/10.1901/jeab.2006.125-04

Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis, 11*(1), 189-196. https://doi.org/10/1901/jaba.1978.11-189

Ingersoll, B. R. (2010). Teaching social communication. *Journal of Positive Behavior Interventions, 12*(1), 33-43. http://dx.doi.org.portal.lib.fit.edu/10.1177/1098300709334797

Johnson, K. (2015). Behavioral education in the 21st century. *Journal of Organizational Behavior Management, 35,* (1-2), 135-150. http://doi.org/10.1080/01608061.2015.1036152

Karsten, A. M. & Carr, J. E. (2009). The effects of differential reinforcement of unprompted responding on skill acquisition of children with autism. *Journal of Applied Behavior Analysis, 42*(2), 327-334. http://doi.org/10.1901/jaba.2009.42-327

Keintz, K. S., Miguel, C. F., Kao, B., & Finn, H. E. (2011). Using conditional discrimination training to produce emergent relations between coins and their values in children with autism. *Journal of Applied Behavior Analysis*, *44*(4), 909–913. http://doi.org/10.1901/jaba.2011.44-909

Kim, J., Wigram, T., & Gold, C. (2008). The effects of improvisational music therapy on joint attention behaviours in autistic children: A randomized controlled study. *Journal of Autism and Developmental Disorders, 38*, 1758-1766. https://doi.org/10/1007/s10803-008-0566-6

Kobari-Wright, V.V., & Miguel, C.F. (2014). The effects of listener training on the emergence of categorization and speaker behavior in children with autism. *Journal of Applied Behavior Analysis, 47*, 431-436. http://doi.org/10.1002/jaba.115

Koegel, R. L., Covert, A. (1972). The relationship of self-stimulation to learning in autistic children. *Journal of Applied Behavior Analysis, 5*(4), 281-387. https://doi.org/10/1901/jaba.1972.5-381

Lanovaz, M. J., Sladeczek I. E., & Rapp, J. T. (2011). Effects of music on vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 3*(44), 647-651. http://doi.org/10/1901/jaba.2011.44-647

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith T. R., & Carr, J. E. (2003). *Behavioral Interventions, 18*(4), 279-289. https://doi.org/10.1002/bin.144

Leblanc, M., Ricciardi, J., & Luiselli, J. (2005). Improving Discrete Trial Instruction by Paraprofessional Staff Through an Abbreviated Performance Feedback Intervention. *Education and Treatment of Children, 28*(1), 76-82.

Lehmann, A. C., & Ericsson, K. A. (1997). Research on expert performance and deliberate practice: Implications for the education of amateur musicians and music students. *Psychomusicology: Music, Mind and Brain, 16*(1-2), 40-58. http://doi.org/10.1037/h0094068

Lim, H. A., & Draper, E. (2011). The effects of music therapy incorporated with applied behavior analysis verbal behavior approach for children with autism spectrum disorders. *Journal of Music Therapy, 48*(4), 532-550. https://doi.org/10.1093/jmt/48.4.532

Lovaas, O. I., Litrownik, A., Mann, R. (1971). Response latencies to auditory stimuli in autistic children engaged in self-stimulatory behavior. *Behavioral Research and Therapy, 9*(1), 39-49. https://doi.org/10/1016/0005-7967(71)90035-0

Mackay, H. A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis and Intervention in Developmental Disabilities, 6*, 373-387. https://doi.org/10/1016/0270-4684(85)90006-0

Makrygianni, M. K., Gena, A., Katoudi, S., & Galanis, P. (2018). The effectiveness of applied behavior analytic interventions for children with autism spectrum disorder: A meta-analytic study. *Research in Autism Spectrum Disorders, 51*, 18-31. http://doi.org/10.1016/j.rasd.2018.03.006

Martens, B. K., Daly, E, J., III, & Ardoin, S. P. (2015). Applications of applied behavior analysis to school-based instructional intervention. In H. S.

Roane, J. E. Ringdahl & T. S. Falcomata (Eds.), *Clinical and organizational applications of applied behavior analysis; clinical and organizational applications of applied behavior analysis* (pp. 125-150, Chapter xv, 661 Pages) Elsevier Academic Press, San Diego, CA. http://doi.org/10.1016/B978-0-12-420249-8.00006-X

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2014). Behavior analysis for lasting change. Cornwall-on-Hudson, NY: Sloan Publishing LLC.

Miguel, C. F. (2018). Problem-Solving, Bidirectional Naming, and the Development of Verbal Repertoires. *Behavior Analysis: Research and Practice*, Advance online publication. http://dx.doi.org/10.1037/bar0000110

Miguel, C.F., Frampton, S.A., Lantaya, C.A., LaFrance, D.L., Quah, K., Meyer, C.S., Elias, N.A., & Fernand, J.K. (2015). The effects of tact training on the development of analogical reasoning. *Journal of the Experimental Analysis of Behavior, 104*, 96-118. http://doi.org/10.1002/jeab.167

Miguel, C.F., Yang, H.G., Finn, H.E., & Ahearn, W.H. (2009). Established derived textual control in activity schedules with children with autism. *Journal of Applied Behavior Analysis, 42,* 703-709. https://doi.org/10/1901/jaba.2009.42-703

Moreno, S., & Farzan, F. (2015). Music training and inhibitory control: a multidimensional model. *Annals of the New York Academy of Sciences, 1337*(1), 147-152. https://doi.org/10.1111/nyas.12674

Motor disorders. (2017). In A. Wenzel (Ed.), T*he SAGE encyclopedia of abnormal and clinical psychology* (pp. 2152-2153). Thousand Oaks, CA: SAGE.

Ogletree, B. T., & Oren, T. (2001). Application of ABA principles to general communication instruction. *Focus on Autism and Other Developmental Disabilities, 16*(2), 102. https://doi.org/10.1177/108835760101600206

Palmer, W.A., Manus, M., & Lethco, A.V. (1981). *Alfred's basic piano library.* Alfred Publishing Company. Partington, J.W., & Sundberg, M.L. (1998). *Assessment of basic language and learning skills (The ABLLS): An assessment for language delayed students.* Pleasant Hill, CA: Behavior Analysts, Inc.

Perkins, D.R., Dougher, M.J., & Greenway D.E. (2007). Contextual control by function and form of transfer of functions. *Journal of the Experimental Analysis of Behavior, 88*(1), 87-102. https://doi.org/10.1901/jeab.2007.36-04

Perez, W. F. & de Rose, J. C. (2010). Recombinative generalization: An exploratory study in musical reading. *The Analysis of Verbal Behavior, 29,* 451-469. https://doi.org/10.1007/BF03393082

Pytte, C. L., & Fienup, D. M. (2012). Using equivalence-based instruction to increase efficiency in teaching neuroanatomy. *Journal of Undergraduate Neuroscience Education, 10*(2), A125-A131.

Rehfeldt, R. A. (2011) Toward a technology of derived stimulus relations: An analysis of articles published in the journal of applied behavior analysis, 1992-2009. *Journal of Applied Behavior Analysis, 44*(1), 109-19. http://doi.org/10.1901/jaba.2011.44-109

Ribeiro, D.M., Miguel, C.F., & Goyos, A.C. (2015). The effects of listener training on the discriminative control by elements of compound stimuli in children with disabilities. *Journal of the Experimental Analysis of Behavior, 104,* 48-62. http://doi.org/10.1002/jeab.161

Santos, P.M., Ma, M.L., Miguel, C.M. (2015). Training intraverbal naming to establish matching-to-sample performances. *Analysis of Verbal Behavior, 31*(2): 162-182. http://doi.org/10.1007/s40616-015-0040-4

Saunders, R.R., & Green, G., (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*(1), 117-137. http://doi.org/10.1901/jeab.1999.72-117

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). The effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences, 1060*(1), 219-230. http://doi.org/10.1196/annals.1360.015

Skinner, B. F. (1959). Cumulative record. East Norwalk, CT, US: Appleton-Century-Crofts

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22. http://doi.org/10/1901/jeab.1982.37-5

Sidman, M., Willson-Morris, M., and Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*(1-2), 1-19. http://doi.org/10/1016/0270-4684(86)90003-0

Stewart, I., McElwee, J., Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The Analysis of Verbal Behavior, 29*(1)*,* 137-155. https://doi.org/10.1007/BF03393131

Stokes T. F., & Baer D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *2*(10), 349-367. http://doi.org/10.1901/jaba.1977.10-349

Suits, B. H. (n.d.) Physics of music – notes. Retrieved from http://www.phy.mtu.edu/~suits/notefreqs.html

Sundberg, M.L. (2008). *VB-MAPP Verbal Behavior Milestones Assessment and Placement Program*. Concord, CA: AVB Press.

Suzuki, S. (1978). *Suzuki Piano School,* Miami, Florida: Summy-Birchard Inc.

Touchette, P. E. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior, 15*, 347-354. http://doi.org/10/1901/jeab.1971.15-347

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review, 30*(4), 387-399. http://dx.doi.org/10.1016/j.cpr.2010.01.008

Vuran, S. (2008). Empowering leisure skills in adults with autism: An experimental investigation through the most to least prompting procedure. *International Journal of Special Education, 23*(1), 174-181.

Watkins, C. L., & Slocum, T. A. (2003). The components of direct instruction. *Journal of Direct Instruction, 3*(2), 75-110.

Whipple J. (2004). Music in intervention for children and adolescents with autism: A meta-analysis. *Journal of Music Therapy*, *2*, 90-106. http://doi.org/10.1093/jmt/41.2.90

Yanardag, M., Akmanoglu, N., & Yilmaz, I. (2013). The effectiveness of video prompting on teaching aquatic play skills for children with autism. *Disability and Rehabilitation: An International, Multidisciplinary Journal, 35*(1), 47-56. http://dx.doi.org.portal.lib.fit.edu/10.3109/09638288.2012.687

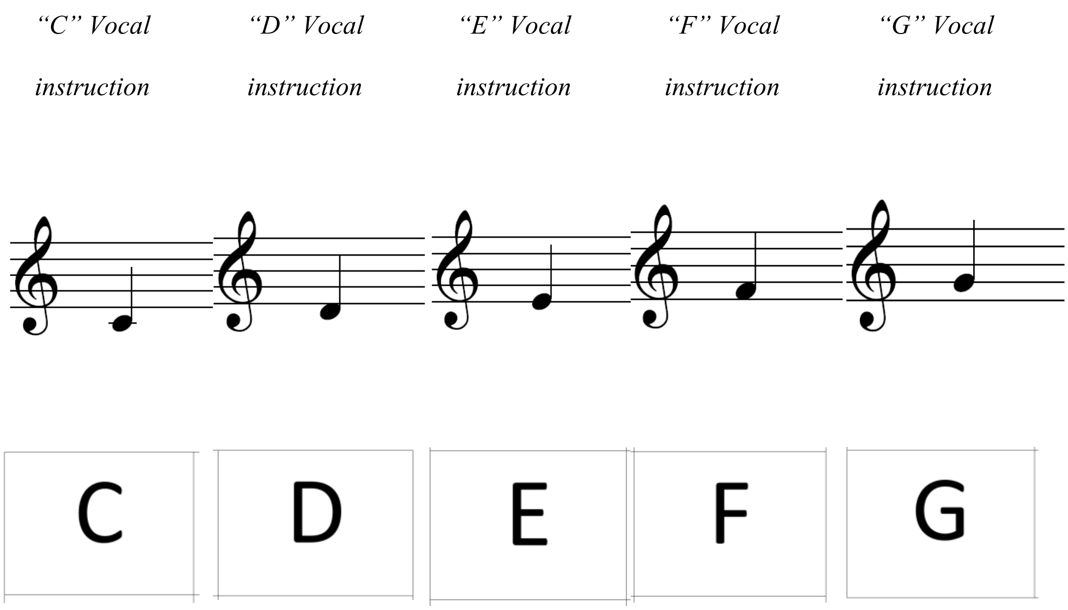

Figure 1. Experimental stimuli.

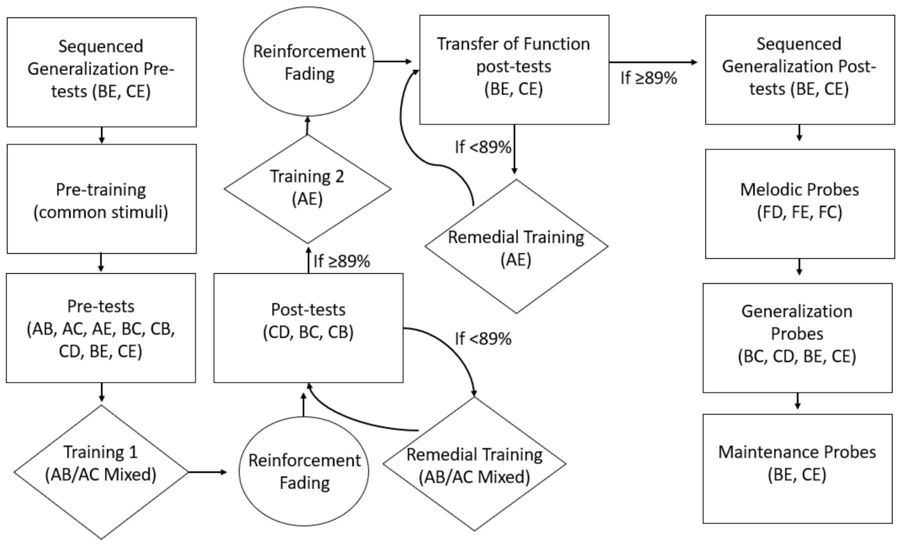

Figure 2. Flowchart of experimental procedures

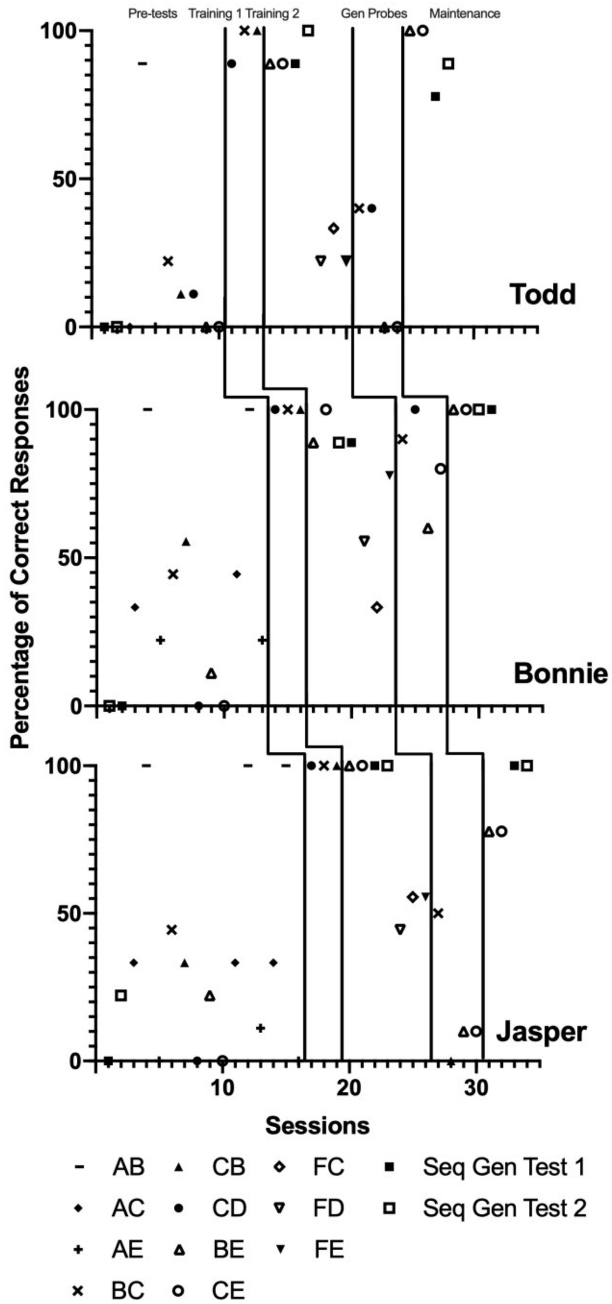

Figure 3. Participant performance for all relations, melodic probes, generalization probes, and maintenance measures.

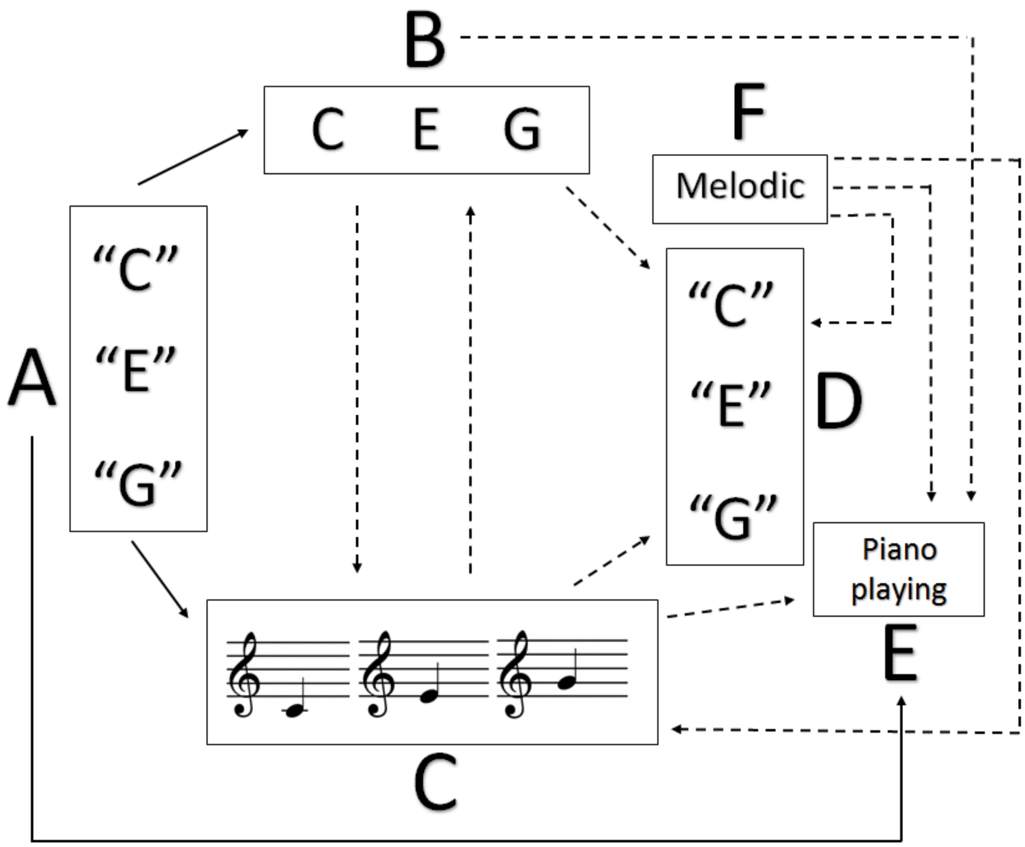

Figure 4. Relations to be trained (solid lines) and tested (dashed lines).

Table 1
Song counterbalancing

Participant	First Song Presented	Second Song Presented
Todd	Version A: Letter	Version B: Musical Notation
Bonnie	Version B: Musical Notation	Version A: Letter
Jasper	Version B: Letter	Version A: Musical Notation

Appendix A

Sample Datasheet

Date:					Participant ID:				
Session:					Primary / IOA / TI				
Task: Training 1: A-B/A-C Mixed Block 1 (Listener training mix)					Participant Block:				
Reinforcement: 100% Social "good job" or "that's right" Error Correction: "No" + CR "this one"					Notes:				
							Tx Integrity		
Trial	Prompt Level	Sample	Left Comp	Middle Comp	Right Comp	Response (+/-/P)	sD	P	SR+
1		G-B	E	G	C				
2		E-C	G	C	E				
3		C-B	C	E	G				
4		E-C	E	G	C				
5		C-B	G	C	E				
6		G-C	C	E	G				
7		C-B	E	G	C				
8		G-C	G	C	E				
9		E-B	C	E	G				
							% Correct:		

Appendix B

Sequenced Generalization Tests BE and CE (versions 1 and 2)

C E C G E G C G E

E C G E C C E G G