



Minnesota State University, Mankato  
Cornerstone: A Collection of Scholarly  
and Creative Works for Minnesota  
State University, Mankato

---

All Graduate Theses, Dissertations, and Other  
Capstone Projects

Graduate Theses, Dissertations, and Other  
Capstone Projects

---

2011

## Utilizing Music in Speech and Language Therapy for Preschool Children and Children with Autism: A Systematic Review

Joelle Kay Johanson  
*Minnesota State University, Mankato*

Follow this and additional works at: <https://cornerstone.lib.mnsu.edu/etds>



Part of the [Speech and Hearing Science Commons](#)

---

### Recommended Citation

Johanson, J. K. (2011). Utilizing music in speech and language therapy for preschool children and children with autism: A systematic review. [Master's thesis, Minnesota State University, Mankato]. Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. <https://cornerstone.lib.mnsu.edu/etds/42/>

This Thesis is brought to you for free and open access by the Graduate Theses, Dissertations, and Other Capstone Projects at Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. It has been accepted for inclusion in All Graduate Theses, Dissertations, and Other Capstone Projects by an authorized administrator of Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato.

Utilizing Music in Speech and Language Therapy for Preschool Children and  
Children with Autism: A Systematic Review

By

Joelle K. Johanson

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

In

Communication Disorders

Minnesota State University, Mankato

Mankato, Minnesota

May, 2011

Utilizing Music in Speech and Language Therapy for Preschool Children and  
Children with Autism: A Systematic Review

Joelle K. Johanson

This thesis has been examined and approved by the following members of the  
thesis committee.

Patricia Hargrove, Ph.D., CCC-SLP, Advisor

Bonnie Lund, Ph.D., CCC-SLP

Karen Hurlbutt, Ph.D.

**Utilizing Music in Speech and Language Therapy for Preschool Children and Children with Autism: A Systematic Review.** Johanson, Joelle K., M.S.

*Minnesota State University, Mankato, 2011. 65pp.*

This thesis was conducted as a systematic review. A systematic review is an assessment and evaluation of current research that attempts to answer a clinical question. The purpose of this systematic review was to attempt to answer the following clinical question: Are there high-quality studies that document the effectiveness of using music in speech and language therapy with preschool aged children and children with autism? This review also attempted to identify future research implications and needs. Nine studies were reviewed and critiqued; six preschool based studies and three autism based studies. Results indicated a general positive change in speech and language outcomes when music is incorporated into speech and language therapy. However, future research needs to be conducted by professionals in the communication disorders field using well designed studies and relevant outcomes to ensure evidence based practice is used among practicing clinicians.

## TABLE OF CONTENTS

### Chapter

I. INTRODUCTION.....	1
Background.....	2
Purpose.....	3
Definitions.....	4
II. METHOD.....	6
Search.....	6
Inclusion/Exclusion Criteria.....	7
Search Results.....	7
Criteria for Review.....	8
III. REVIEWS.....	10
Preschool Based Studies.....	10
Aldridge, Gustroff, & Nuegebauer (1995).....	10
Gross, Linden, & Ostermann (2010).....	14
Hoskins (1998).....	16
Kouri & Winn (2006).....	19
Ross (1997).....	23
Seeman (2008).....	26
Autism Based Studies.....	28
Edgerton (1994).....	29
Lim (2010).....	32
O'Loughlin (2000).....	36

IV. DISCUSSION.....	40
Preschool Based Studies.....	40
Autism Based Studies.....	43
Clinical Implications.....	45
Future Research.....	49
References.....	51
Appendix A.....	55
Appendix B.....	60
Appendix C.....	61
Appendix D.....	64

## LIST OF TABLES

1. Reviewer's grading criteria.....	9
2. Reviewer's overall grading for preschool based studies.....	41
3. Reviewer's overall grading for autism based studies.....	44
4. Clinical implications of preschool based results.....	47
5. Clinical implications of preschool based results.....	48

## Chapter I

### Introduction

In any healthcare profession that provides services to clients, it is imperative to use methods that are shown to have effectiveness and efficacy. The field of speech-language pathology has recently established guidelines for speech-language pathologists (SLPs) to follow that ensures effective methods are used when conducting assessments and treatment, and evidence-based practice (EBP) is being applied. Practicing clinicians need to be able to apply EBP skills to their everyday decision making to ensure they are providing the best possible evaluation and treatment methods for their clients.

The American Speech-Language-Hearing Association (ASHA) has adopted a broad definition of the term EBP defined as “the integration of best research evidence with clinical expertise and patient values” (Sackett, 2000, p. 1). According to ASHA’s website, “the goal of EBP is the integration of: (a) clinical expertise, (b) best current evidence, and (c) client values to provide high-quality services reflecting the interests, values, needs, and choices of the individuals we serve” (ASHA). When making clinical decisions that are EBP based, SLPs must:

evaluate the efficacy, effectiveness, and efficiency of clinical protocols for prevention, treatment, and enhancement using criteria recognized in the evidence-based practice literature [and] evaluate the quality of evidence appearing in any source or format, including journal articles, textbooks,



continuing education offerings, newsletters, advertising, and Web-based products, prior to incorporating such evidence into clinical decision making. (ASHA, 2005)

## **Background**

The use of music within speech and language therapy has been growing within the field of speech-language pathology. The use of music in speech and language therapy has been applied with many populations ranging from infants to adults. Populations in which music therapy have been used that were relevant to the field of speech-language pathology have included articulation disorders, language disorders, apraxia, cochlear implant clients, clients with aphasia, and clients using augmentative and alternative communication (AAC).

This systematic review focuses on research concerning the effectiveness of using music within speech and language therapy with preschool children and children with autism. Zoller (1991) stated, “actively using music in learning experiences involves the whole child through incorporation of rhythm, movement, and speech. Within the public school setting, traditional communication training methods can be supplemented with musical activities” (p. 272).

Presentations have been conducted at ASHA conventions throughout the years regarding the practice of incorporating music. Arntson (2006) presented on the benefits of using music in therapy stating it involves active participation of the client, the use of memory, motor imitation, emotion, and provides repetition for additional practice. In 2009, Arntson also presented on music and autism at the

ASHA convention and claimed music can help facilitate success within this population. Arntson also pointed out songs can provide predictability and a cueing system that can be gradually decreased.

Collaboration (i.e., to work with) and consultation (i.e., to ask for advice or meet with) between music therapists and SLPs have been an increasing occurrence in the communication disorders field. Boucher (2008) presented on the collaboration of music therapists and SLPs stating “music therapists work in collaboration with speech language pathologists in incorporating music into nonmusic goals to best meet the needs of a group or individual.” McCarthy and Geist (2007) conducted a survey regarding the collaboration between the two professions. Results indicated that 36.3% of music therapists reported consulting with SLPs, and 44.3% of music therapists reported collaborating with SLPs. The benefits of collaboration according to this survey included enhancing knowledge about music therapy with SLPs, enhancing goals, enhancing client progress, enhancing professional support, and enhancing ingenuity. The survey did not include any information from SLPs and their collaborated work with music therapists.

### **Purpose**

The previously mentioned information and several other published materials provide some support that using music in speech and language therapy is beneficial but there are concerns with the quality of the empirical support by using music in speech and language therapy. The purpose of this thesis is to

review systematically the available evidence utilizing music in speech and language therapy with preschool children and children with autism. The clinical question this review attempts to answer is: Are there high-quality studies that document the effectiveness of using music in speech and language therapy with preschool aged children and children with autism? This review will also attempt to identify future research implications and needs.

### **Definitions**

Articulation disorder – “the atypical production of speech sounds characterized by substitutions, omissions, additions or distortions that may interfere with intelligibility” (ASHA, 1993)

At-risk – children who “lack early experiences that support their growth and development” and therefore are susceptible to a developmental delay (Minnesota Department of Education, 2009, p. 3)

Autism – a neurodevelopment condition with a neurological base representing a spectrum of difficulties in socialization, communication, and behavior (Paul, 2007)

Developmental delay/mental retardation – “a disability characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed by conceptual, social, and practical adaptive skills” (Paul, 2007, p. 108)

Evidence based practice – “the integration of best research evidence with clinical expertise and patient values” (Sackett, 2000, p. 1)

Language disorder – impairment in “comprehension and/or use of a spoken, written, and/or other symbol system. The disorder may involve (1) the form of language (phonologic, morphologic, and syntactic systems), (2) the content of language (semantic system), and/or (3) the function of language in communication (pragmatic system), in any combination” (ASHA, 1993)

Main effect – “the individual effect of each independent variable and each parameter of the dependent variable” (Schiavetti & Metz, 2006, p. 210)

Music therapy – “an established, accredited health profession whereby the systematic application of music is utilized in the treatment of cognitive, social, communicative, behavioral, psychological, sensory-motor, and physical needs of an individual” (Boucher, 2008)

Nordoff Robbins approach – an approach using music in therapy with children and adults who live with a mental and/or physical disability, neurological damage, Down syndrome and other causes of developmental delay, autism spectrum disorder, and several other populations (Nordoff Robbins, 2011)

Significant difference – “the degree of confidence that the researcher has that the difference seen in the sample data would not have occurred by chance alone” (Schiavetti & Metz, 2006, p. 185)

## Chapter II

### **Method**

A systematic review is an assessment and evaluation of research that attempts to answer a clinical question. Systematic reviews present an organized literature review resulting in a conclusion that “can be made from a larger group of studies that cannot be made based on individual studies alone” (McCauley & Hargrove, 2004, p. 174). The method chapter will discuss the search process, inclusion and exclusion criteria for the studies, the final search results, and the criteria for reviewing the studies. Some of the major factors that differentiate systematic reviews from traditional, or narrative reviews of the literature are: (a) the strategy for searching the literature is described, (b) inclusion and exclusion criteria for selection of sources is established prior to the search, (c) prior to the review criteria for analyses are agreed upon, (d) all sources are subjected to analysis using the analysis criteria.

### **Search**

Computer searches were conducted to retrieve appropriate studies for this review. The following databases were searched: RILM Abstracts of Music Literature, ERIC, ComDisDome, Google Scholar, Medline, psycInfo, CINAHL, Masterfile Premier, and Professional Development Collection. Reference lists from retrieved studies were also examined to identify any articles that might have been appropriate for this review. The terms used to search these databases

included speech therapy AND music, language therapy AND music, speech-language pathology AND music, voice AND intervention OR therapy OR acquisition AND music, aphasia AND intervention OR therapy OR acquisition AND music, fluency AND intervention OR therapy, speech OR articulation OR language OR phonological awareness AND intervention OR therapy OR acquisition AND music.

### **Inclusion/Exclusion Criteria**

The following inclusion and exclusion criteria were developed to review studies that contained areas of interest to the reviewer. Studies that were reviewed met the following criteria: (1) articles were not a description of a curriculum, (2) participants were at the preschool level with an age of less than 5 years old or school aged children with autism, (3) articles were not reviews of the literature, and (4) studies were not case studies. There was no year limit to the studies.

### **Search Results**

The initial computer search produced 836,521 possible articles. Articles were first eliminated based on titles which narrowed the search down to 38 articles. From there, abstracts or introductions were read to ensure articles met the inclusion criteria. This analysis yielded 19 studies. Of these 19 studies, nine publications met all the criteria. These nine publications were then divided into two categories: preschool based studies and autism based studies.

## **Criteria for Review**

**Summaries.** Each of the nine articles were summarized by the reviewer. An initial grade was given based on the type of evidence described in the study. The investigator analyzed each article using the form in Appendix A. This was important to determine reliability, efficacy, and generalization of results. The summarization included an initial grade for the type of evidence identified (see Appendix B for grading criteria), the purpose of the study, the method used, and the results. Attention was paid to the number of participants in each study, any defining variables such as age or level of impairment, the method of intervention the author(s) chose, and how the intervention was conducted.

**Critiques.** A critique was completed for each of the nine studies by the investigator and her advisor. The investigator and her advisor independently analyzed each study using Appendix A. The investigator and advisor then discussed each study and came to a consensus on any disagreements. The critique was conducted following the same Appendix A form. The critique focused on the type of evidence the study presented, participant and group information (if applicable), outcomes, statistical results, and any evidence of EBP measures. Following each critique the reviewer assigned a final overall grade to each study based on the assessment of the quality of evidence. Overall grading was based on the holistic view of the design and quality of each study. See Table 1 for reviewer's overall grading implications.

Table 1

*Reviewer's Grading Criteria*

---

Grade	Description
A	High quality of evidence, well designed study, large sample number, well described and generalizable therapy technique, relevant outcomes, EBP metrics used in analysis. Would use this technique in the clinical setting.
B	High quality of evidence, well designed study, adequate sample size, generalizable therapy technique, relevant outcomes, EBP metrics used in analysis. Would consider using this technique in the clinical setting.
C	Moderate level of evidence, adequate design of the study, moderate sample size, described technique, relevant outcomes, analysis was done on data. Would contemplate whether this approach would be ethical to use in the clinical setting.
D	Low quality of evidence, poor study design, low sample size, described technique, relevant outcomes, analysis were done. May draw specific techniques to use in the clinical setting based on relevant outcomes.

---



## Chapter III

### Reviews

This chapter will present each preschool based article's review and critique followed by each autism based article's review and critique. The following reviews and critiques were conducted in a stylized manner ensuring specific topics (i.e., purpose of the study, the population, the method used, and the results) were analyzed.

#### Preschool Based Studies

The search provided six studies of intervention designed to treat preschool children using music in speech and language therapy. The following section contains a detailed summary and critique of each preschool-based study.

##### **Aldridge, Gustroff, & Neugebauer (1995).**

**Summary.** The purpose of this study was to determine if music therapy would have a positive effect on developmental changes. The design of this study was prospective, randomized group design with controls which has an assigned A for level of support. The design initially had a treatment and a no-treatment group as a waiting-list control group. The non-treatment group received therapy after waiting for three months and the initial treatment group had a three month period without therapy. The design followed a course of an ABAB treatment, alternating between music therapy and no treatment with periods lasting three months. Assessments were conducted every three months following the treatment switches. The music therapy was developed from the Nordoff and

Robbins approach. Assessments were conducted using the *Griffiths* scale designed to diagnose areas of a child's capability and to provide a profile for treatment. The *Griffiths* scale has six subscales. Results will be reported in this review on the three of the six subtests which were related to communication disorders: personal-social scale, hearing and speech, and practical reasoning.

There were a total of eight participants in two groups. The initial therapy group contained five participants and the initial control group contained three participants. The authors stated the participants were randomly assigned to the two groups. Group membership was not concealed from the participants or the clinicians. The analyzers were initially blind to group membership. The authors stated the participants had to have a chronological age of 4.0-6.5 years and a developmental age of 1.5-3.5 years in order to qualify for the study. Exclusion criteria included previous music therapy, degenerative physical problem, any psychopharmaceutical treatment, and previous creative art therapy. The authors did not include any information regarding gender, expressive language skills, receptive language skills, mean length of utterance, socioeconomic status, and educational level of clients or parents. The authors reported at the beginning of intervention the two groups were similar in chronological age but different in mental age. The authors maintained 83% of original members in group one and 50% in group two. The initial test period was the only time when the study compared a treatment and control group.

Outcomes were reported based upon the subscales of *Griffiths* test. The outcomes from this test were subjective. The authors did not report any reliability measures. The outcome measurements hearing and speech and personal-social subtest were significant for both treatment groups. The practical reasoning subtest did not have significant outcomes. Outcomes were reported with a p value for every administration period. Test 1 will be the only reported outcome in this review because it was the only time an experimental and control group could be compared. The probability levels for hearing and speech were  $p=0.004$ , probability levels for personal-social were  $p=0.044$ , and probability levels for practical reasoning were  $p=0.188$ .

**Critique.** The treatment followed the Nordoff and Robbins approach. The authors did not include specific information regarding the treatment process, making it difficult to recreate the study to apply this approach clinically. There was also little information regarding the validity of the *Griffiths* scale. The authors described the scale and stated the developmental measure was based on the child's mental age. There was little information on how the individual participants' profiles reflected their functional level.

The members of this study were not adequately described. The only information the authors provided were inclusion and exclusion criteria. Information regarding cognitive and linguistic information could only be interpreted from what was provided through the results of the *Griffiths* tests. This makes the treatment approach difficult to generalize in the clinical setting.

Participants were lost during the experiment. Group one maintained 83% of its participants and group two maintained 50% of its participants. A final total of eight participants does not allow for generalization of this approach.

Results from the study were reported in terms of changes between testing periods on the *Griffiths* test. It is unclear whether this test is an appropriate measure to determine effectiveness. Change between the testing periods did reveal significantly effective changes. The outcome measurements hearing and speech and personal-social subtest were significant. The practical reasoning subtest did not have significant outcomes. Outcomes were reported with a p value for every administration period. The practical reasoning subtest, which the authors reported was the most dependent on speech and represents the general language of participants, showed no significant effect. The accuracy of the p values is questionable because the sample sizes are too small for the type of analysis conducted. A nonparametric analysis would have been more appropriate with this sample size. There were no evidence based measures provided.

The overall grade for this study was a D. The overall design of this study was well developed. However, it is unknown if the *Griffiths* test has any validity in the communication disorders field. The outcomes for two measurements were significant; although the subtest the author claimed to represent speech and language the most had no significant effects.

**Gross, Linden, & Ostermann (2010).**

**Summary.** The purpose of this study was to explore the effects of music therapy on verbal reasoning abilities in children with delayed speech development. There were a total of 18 participants between the ages of 3.5 and 6 years old with delayed speech development. This pilot study was a prospective, single group with pre- and post-testing with a level of evidence grade of B-. Therapy was designed as an ABAB treatment, alternating between music therapy and no treatment with periods lasting eight weeks. Assessments were administered before and after each study period.

Therapy service was provided through an out-patient basis at the Department of Music Therapy at a community hospital. Sessions were conducted by two music therapists. Assessments between each treatment period were conducted by a speech-language pathologist and psychologist. Outcomes measured included cognition and speech development; cognitive development will not be discussed in this review. Specific outcomes related to speech development included phonological memory for nonwords, memory for sentences, generation of morphological rules, and memory for word sequences.

Therapy methods for this study were based on the Nordoff Robbins approach. In this approach, patient and therapist were active in singing and making music with percussion instruments and a piano. The authors did not provide any further information regarding therapy methods.

The authors concluded that music therapy had a moderate effect on memory for sentences with a standardized mean difference score (d) of .61. There was a small effect on phonological memory for nonwords and understanding sentences with a d value of .45 and .39 respectively. There was no treatment effect on the outcomes of generation of morphological rules and memory for word sequences.

**Critique.** The type of evidence identified was prospective, single group with pre- and post-testing. This type of evidence has a level of B-. The study did not include a control group to compare the effectiveness of the treatment method. The participants in the only experimental group were not randomly assigned which affects the generalization of this treatment strategy. Group membership was not concealed from participants or clinicians although the authors attempted to conceal membership from the analyzers. However, because these analyzers interacted with the children and saw them five times throughout the course of the study, concealment could not be concluded.

The members of this study were not adequately described. A total of 18 participants in the one experimental group is not a large sample size which minimizes the ability to generalize. Ages were between 3.5 to 6 years and included 6 females and 12 males. All participants had to have a developmental speech disorder and had to score below a score of 50 in a subtest of short-term memory for non-words on a formal assessment. Further information regarding description of participants was minimal. The authors did not include any

information about cognitive status or specific information concerning expressive and receptive language ability. The authors did not include a sufficient amount of information to recreate study or to apply this method clinically. On the positive side, the treatment group maintained at least 80% of their original members throughout the study.

The Friedman test was used to analyze the effect of music therapy over the course of time and the Wilcoxon rank-sum test was used for baseline comparisons and final measurements after the last therapeutic session. The authors reported an effect size using standardized mean difference. Three of the five outcomes measuring language had an effect. Memory for sentences had a moderate effect ( $d=.61$ ). Phonological memory for nonwords ( $d=.45$ ) and understanding sentences ( $d=.39$ ) both had small effects.

Overall quality of this study is a C due to the authors not providing enough information regarding participants and treatment method. However, the treatment method did make small to moderate effects on three outcomes.

### **Hoskins (1988).**

**Summary.** The purpose of this study was to investigate if sung versus spoken administration of standardized tests would show a relationship with responses and to investigate the use of music activities to increase expressive language abilities of language delayed preschoolers. There were three groups based on ability level within this study. The three groups were constant in time and treatment strategy. The design of Hoskins research cannot be strictly viewed

as clinical research. The design that the present research most closely resembles is a single subject with pre- and post-tests. Accordingly, this investigator assigned Hoskins a grade of B-. There were two pretests administered prior to intervention. The *Expressive One-Word Picture Vocabulary Test (EOWPVT)* and the *Peabody Picture Vocabulary Test (PPVT)* were both administered according the test manual. The *PPVT* was then administered in a melodic version. Following the pretests, therapy sessions were initiated and were conducted three days per week for 30 minute sessions. Participants were assigned to three separate groups based on chronological age and functional abilities. Sessions consisted of music activities with emphasis on increasing expressive language skills. A picture of an object was shown to the group and the therapist sang a three to five word phrase about the object. The group then repeated the name of the object with the therapist. Following the intervention period, posttests identical to the pretests for language ability were administered.

There were a total of 16 participants assigned to three groups which were established according to chronological age and functional abilities of the children. Functional ability levels were high functioning, moderate functioning, and low functioning. The ages of the participants were 2-5 years. There were eight males and eight females. The participants' IQ was reported ranging between 44 and 100 with a mean of 74. The participants had no hearing difficulty and were capable of some speech. All groups received intervention. The three groups maintained at least 80% of their original members throughout the study.



The outcomes measured were the *PPVT* both spoken and melodic version and the *EOWPVT*. The *PPVT* spoken version had no main effect for trials and group by trials. The author conducted the follow up Scheffe test which indicated the high ability participants were significantly different ( $p < .05$ ) from the low ability participants. The *PPVT* melodic version indicated that the high ability group was significantly different ( $p < .05$ ) from the other two groups. Significant improvement was found for all participants ( $p < .05$ ) with no significant group by trials interaction on the *PPVT* melodic version. Wilcoxon analysis of the total sample was conducted for the pretest and posttest which showed a significant improvement ( $p < .05$ ) for the melodic *PPVT*. The author also conducted the Walsh test which indicated the moderate group ability showed a significant improvement ( $p < .05$ ) for the melodic *PPVT*. The *EOWPVT* results indicated the high ability group was significantly different ( $p < .05$ ) than the other two groups. There were no evidence based practice measures.

**Critique.** The design of this experiment was detailed and thoroughly executed. All participants received the same pretests, same intervention, and the same posttests. Using the same tests for the pretest and posttest showed if there was improvement but the reliability has to be viewed cautiously after administering the same test twice within a short time period. There were three independent groups within this study so it did not completely fit the prospective, single group with pre- and post-testing design. This specific design may be

viewed as slightly stronger evidence because of the ability to compare between the three groups.

The author did not provide enough information about the music activities during the treatment phase to use this method clinically. The administration of the standardized test in a melodic version did not have any clinical appeal to this reviewer; it would not be feasible for speech-language pathologists to administer tests in this way. The author did not provide a sufficient amount of information regarding the participants in this study which has a negative impact on generalization.

The outcomes of this study showed improvement on both the *PPVT* (more effect in the melodic version) and *EOWPVT*. It is difficult to determine if this is due to the re-administration of the same test within ten weeks or if the intervention was effective. The overall grade for this study was a C-.

**Kouri & Winn (2006).**

**Summary.** The purpose of this study was to examine how singing affects children's quick incidental learning (QUIL) of vocabulary terms. QUIL is defined as a child's ability to learn a new word on the basis of just a few exposures to it in order to rapidly expand their vocabularies. Participants were presented with spoken and sung story scripts containing novel words over two experimental sessions to determine if preschoolers with mild developmental delay and specific language impairment were able to acquire novel lexical terms and if children's comprehension or production of these novel terms varied as a function of

exposure to sung versus spoken script. The experiment was a prospective, single group with pre- and post testing with a level of support B-. Two experimental sessions lasting approximately 50-60 minutes in length were conducted where participants were randomly assigned to groups hearing both sung and spoken scripts. The two scripts were counterbalanced across participants and sessions. Results regarding the spoken scripts will not be addressed in this summary and critique. Testing was conducted before the initial session and after the two experimental sessions.

Participants were not randomly assigned to the single experimental group and concealment of participation was not achieved. The participants had to have a delay of at least 12 months or 1.5 standard deviations below the mean on one or more standardized tests. There were a total of 16 participants with ages ranging from 3.6-5.1 years. Exclusion criteria included the presence of a hearing loss or any neurological disorders. Overall language skills had to be 1.5 standard deviations below the mean on at least one standardized test. Standardized tests conducted before the experimental sessions began were *Sequenced Inventory of Communication Development*, *Clinical Evaluation of Language Functions*, *Preschool Language Scales*, and *Battelle Developmental Inventory*. The authors did not administer all standardized tests to every participant. An average mean length of utterance was 2.00. A mean score was reported for the cognitive section of a standardized test.

Group membership was maintained throughout the study. Outlying data were not removed from the study. Interobserver reliability was reported at 93% for production of probes and 82% for unsolicited productions. The authors did not include any data regarding intraobserver reliability or treatment fidelity.

A within-subjects analyses of variance (ANOVA) was conducted with the sung condition and session (one and two) as the independent variables. The outcomes measured were subjective values. Outcomes reported were comprehension of novel items, production of these items during the sessions, generalization when presented with distracter foils, and unsolicited productions of novel items. Between experimental session one and two comprehension, production, and unsolicited production of the probes improved. The generalization between session one and two decreased. After two experimental sessions, the authors reported production of unsolicited imitations of novel items in the sung condition with a large effect size of  $d = 0.80$ . In the sung condition, comprehension, production, and generalization were not better. For the unsolicited production, the sung condition was better than spoken. Overall, the sung and spoken conditions combined and number of session revealed a large effect size of  $d = 0.80$  and a significant effect with  $p = 0.05$ .

**Critique.** The type of evidence identified was prospective, single group with pre- and post-testing. This type of evidence has a level of B-. The study did not include a control group to compare with the treatment method effectiveness.

The participants in the only experimental group were not randomly assigned which affects the generalization of this treatment strategy.

The authors included a minimal description of the participants. They displayed a table of participant information acquired from the pretesting portion that detailed the overall experimental population. However, the authors did not include any detailed information regarding gender, cognitive skills, or socioeconomic status. The authors did not administer every test to every child. The mean values on the standardized tests then did not reflect the whole group.

The authors appeared to have included enough information about the treatment method to apply clinically. They provided an appendix which included exact scripts to the stories and songs used during the sessions and questions the clinicians used to probe the targeted responses.

Results reported were comprehension of novel items, production of these items during the sessions, generalization when presented with distracter foils, and unsolicited productions of novel items. Comprehension, production, and unsolicited production improved between the two sessions, and generalization declined. For the unsolicited production of the probes, there was a large effect between the sung condition and number of treatment sessions. Overall, both the sung and spoken had a large and significant effect. Overall grade for this experiment was a C+.

**Ross (1997).**

**Summary.** The purpose of this study was to determine the effect of singing on the articulation of children with language impairments. The author of the study investigated if singing would increase the frequency of the target sounds /M/, /P/ and /B/ (sic) during the music therapy session and in the classroom. This single subject experiment had a multiple baseline design across three participants between the ages of 3 and 5. The level of support for this design study was an A-. The three participants were subject J, subject C, and subject T. Music interventions consisted of 16 30-minute therapy sessions. Data was recorded on the target sounds in the therapy room and the classroom before and after each session.

There were a total of three participants. The characteristics of the participants described were age, gender, cognitive skills, expressive language, and educational level of participants. In order for the participants to have been considered for this study, they had to be diagnosed with a severe expressive speech impairment by a speech-language pathologist. Other impairments included developmental delay and Down syndrome. Membership of the participants was maintained throughout the study.

Baseline data were collected on all behaviors. The probes tested production of /m/, /p/, and /b/ with continuous data collection. The outcome of these targets was measured subjectively. In general, the authors claimed target phonemes improved during and after treatment.

The author did not include statistical data in relation to effectiveness. Using percentage of nonoverlapping data (PND), the overall effectiveness of the targeted phonemes generalized into the classroom was determined. PND is calculated by taking the highest value during baseline then determining the percentage of data points in the treatment time that indicated a better performance than the highest value from the baseline. Subject J achieved 100% PND with an interpretation of highly effective across all three target sounds. Subject C achieved 71% PND for phoneme /m/ with an interpretation of fairly effective, 0% PND for phoneme /p/ with an interpretation of ineffective, and 75% PND for phoneme /b/ with an interpretation of fairly effective. Subject T achieved 100% PND for phoneme /m/ with an interpretation of highly effective, 0% PND for phoneme /p/ with an interpretation of ineffective, and 25% PND for phoneme /b/ with an interpretation of ineffective. Overall, the results suggested that treatment was effective for two of three participants.

**Critique.** The focus of this research was clinical research. The type of evidence identified was single subject experimental design with specific client with multiple baselines. This type of evidence has a level of A-. The study did not include a control group to compare treatment method effectiveness. However, in this design study baseline was considered to be a control. Group membership was not concealed from participants, clinicians, or the analyzers. Data were recorded by a music therapist with a speech therapist as the secondary observer.

Overall, the participants in this study were not adequately described. There was a total of three participants. A total of three participants in a single subject design moderately limits one's ability to generalize the treatment strategy. The ages of the participants were between 3 to 5 years and were chosen by a speech-language pathologist with criteria having to meet a severe expressive speech impairment. Other disorders included developmental delay and Down syndrome. Further information regarding description of participants is minimal. The author did not include any information receptive language ability which would be beneficial to know before and after treatment. The treatment group maintained at least 80% of their original members throughout the study.

The authors did not include a sufficient amount of information to recreate study or to apply this method clinically. The experimental intervention was implemented during regular music therapy sessions. The author stated that songs were sung during intervention and targeted the three sounds /m/, /p/, and /b/ were included in the appendix of the article. The author stated in each session there was a hello song, an instrumental activity, a cognitive activity, the song intervention activity, a group movement activity, and a goodbye song.

The author reported data regarding accuracy of the targeted sounds for each student in both the classroom and intervention settings. However, the author did not provide statistical analysis.

Percent of nonoverlapping data (PND) was calculated by this reviewer to determine effectiveness of this treatment strategy for generalization into the



classroom setting. Overall data suggested that treatment was effective for two of three participants. The most improved articulation sounds were /m/ and /b/.

The overall quality of this study was C- due to the authors not providing enough information regarding participants, treatment method, and evidence based practice measurements.

**Seeman (2008).**

**Summary.** The purpose of this study was to determine short-term effects of music education on receptive language skills of students in an early childhood program in an at-risk community. This was a prospective, single group design with pre- and post-testing which has a level of evidence B-. The students participated in ten weeks of intervention, with two sessions per week. Pre- and post-testing was conducted to evaluate receptive vocabulary, language development ratings using two standardized tests: *Peabody Picture Vocabulary Test (PPVT)* and the *Teacher Rating of Oral Language and Literacy (TROLL)*.

Originally, there was a total of ten participants in the study. One participant was lost during the experiment; original data from that participant was removed. The participants were not randomly assigned to the single experimental group. Participants were selected based on presence of at-risk or special needs and teacher recommendation. Group membership was not concealed from participants, clinicians, or analyzers. The author described the age and gender of the participants, ranging from 3.60-4.10 years old with three females and six males. Information about the race of the participants were included; eight

participants were Caucasian and one was African American. The authors did not include any information regarding cognitive skills, expressive language skills, mean length of utterance, socioeconomic status, and educational level of clients or parents. The group maintained 90% of participants throughout the study.

Reliability measures were not provided by the author. The final outcomes reported were the age equivalents on the *PPVT* and scores on the *TROLL*. The author concluded there was a 21.18% increase on the age equivalent for the *PPVT*; starting from a mean age of 4.6 and ending with 6.1. The author reported an increase in all categories of the *TROLL*. An increase of 43% for communicate personal experiences, a 33% increase for recognize and produce rhymes, and 28% increase for use a varied vocabulary. The author did not provide any evidence based practice measures.

**Critique.** The prospective, single group design has a grade of B-. The author did not include a control group to compare to the experimental group. The participants were selected out of an early childhood program. There was the opportunity to use the remaining children as a no intervention.

The author went into great detail about the intervention. Appendices and schedules were included within the article. The amount of information provided from the author gives the ability to reproduce the experiment for someone who is interested in this treatment approach.

The group was not adequately described. There was not enough information regarding the participants' status for functional level, expressive

language skills, or socioeconomic level which is an indicator for at-risk students. The lack of information negatively impacts generalization of this treatment.

The outcomes the author reported cannot be considered applicable in the field of communication disorders. The author reported age equivalent scores on the *PPVT*, which have little to no clinical application. The *PPVT* has the ability to produce standardized scores which are more relevant to clinicians. The *TROLL* appeared to be a subjective measure of abilities. It is a rating scale for teachers to fill out based on the students' performance. A more objective assessment could have been used to provide further quantitative data. The author did not include any effect outcomes of the treatment condition. The results of the experiment were presented in a narrative form with no evidence based practice measures. From the author's reported scores, we can assume the treatment implemented was successful with increases throughout all variables.

The overall grade for this study was a D-. The author included a sufficient amount of information to apply this method clinically. However, the lack of participant information for generalization, the reported outcomes based on age equivalent scores, and no presence of evidence based practice negatively impacts the efficacy of this approach.

### **Autism Based Studies**

The search provided three studies of intervention designed to treat children with autism using music therapy. The following section contains a detailed summary and critique of each autism based study.

**Edgerton (1994).**

**Summary.** The purpose of this study was to determine if music therapy designed on the Nordoff and Robbins approach will the expressive communication skills of children with autism. The design of this research was single subject experimental design with specific clients with a grade of A-. The author implemented a reversal design after six weeks. The intervention was improvisational music therapy designed from the Nordoff Robbins approach. Each participant attended one 30-minute session for 10 weeks. Throughout the sessions a hierarchy of musical activities was implemented dependent on the child's responses, capacities, and needs. The clinician attempted to establish contact with the child and enable the child to respond to facilitate communication. A checklist designed specifically for this study called the *Checklist of Communicative Responses/Acts Score Sheet (CRASS)* assessed communicative behavior in terms of communicative responses and communicative acts for musical and nonmusical communication. For this review, only the nonmusical results will be reported. The *CRASS* was administered during every session with each child. Following the 10 weeks of intervention, *The Behavior Change Survey* was completed by parents, teachers, and speech therapists.

There was a total of 11 participants diagnosed with autism ranging from mildly impaired to severely impaired. The age of these participants was between 6 and 9 years. There were 10 males and one female. The participants' verbal

skills included 5 nonverbal children and four with limited functional language skills as determined by a speech-language pathologist. Membership during the study maintained at least 80%. One participant attended 80% of the sessions, while another attended 90%.

The *CRASS* and *The Behavior Change Survey* were used to determine effectiveness of this approach. The author included interrater reliability measures for the *CRASS* with an occurrences mean of 86.2% and nonoccurrences mean of 94%.

The author included information regarding musical and nonmusical communication acts. For this review, nonmusical will be reported. A Wilcoxon Matched-Pairs Signed-Ranks Test showed all of the participants' last session scores were significantly greater than their first session scores at the .01 level. Significant differences were found at the .05 level between first and last session scores for communicative intent. *The Behavior Change Survey* was completed by 11 parents, 4 teachers, and 2 speech-language pathologists and indicated that most means fell between 4 which indicated no change and 5 which indicated a slight change. Both communication and social/emotional categories received a mean of 4.5.

**Critique.** The design of this study was difficult to identify. The design was determined to be a weak single subject experimental design with specific clients. The grade assigned to this design is an A-. However, due to the weakness this grade should be viewed with caution.

The participants were not adequately described in this study. The author provided age and gender. The author also included very vague narrative information about the expressive skills regarding only nine of the eleven participants. The functional level of the participants was described between mildly impaired to severely impaired. This gives very little information regarding the participants and does not provide for generalization. There was no control group in this study to compare the effectiveness of this treatment.

The outcomes for this study were subjective measurements. The *CRASS* was designed by the author and it was a checklist of whether or not a behavior was observed during a session. The second outcome measure was a survey completed by the parents, teachers, and speech-language pathologists. This was a rating scale to indicate any changes in communicative, social/emotional or musical behaviors. The outcome of the rating scale reported no change to slight change across participants as rated by speech-language pathologists. Rating scales are seen as highly subjective measurements and can vary between raters. The author did not provide any interrater reliability measurements on *The Behavior Change Survey*. The outcomes showed a stable positive trend among all participants in the study. The author did not provide any evidence based practice measurements to show indication of a positive effect on the participants.

An overall grade for this study was a D+. A sufficient amount of information about the participants was not included which negatively effects the generalization of this method. The author did report a positive change on the

CRASS across all participants. However, the CRASS was developed by the author and has no evidence behind this form of assessment. There was little to no change on *The Behavior Change Survey* for the participants. The author did not provide any evidence based practice measures to show the effectiveness of this method.

**Lim (2010).**

**Summary.** The purpose of this study was to examine the effects of musical patterns on the perception and production of speech in children with autism spectrum disorder. The author investigated if speech production differed by training conditions, level of functioning, and if any interaction exists between training condition and aspects of speech production: semantics, phonology, pragmatics, and prosody. This was a prospective, randomized group design with controls which has a grade of an A for level of support. The design involved three groups: music condition, speech condition, and no treatment group. The music therapy in this study was developed using a training called developmental speech and language training through music (DSLTM). Songs were created containing 36 target words paired with pictures presented for the target words using The Picture Exchange Communication System (PECS). The speech condition included the same texts for the songs used in the music stimuli along with the presented pictures. Participants were all given a pretest and then completed six training sessions of the music or speech condition, followed by

administration of the posttest. The pretest and posttest, the *Verbal Production Evaluation Scale (VPES)*, was developed by the investigator.

There were 50 participants in this study. The participants were randomly assigned to one of three conditions: music, speech, or control group. The music and speech condition both contained 18 participants and the control group contained 14 participants. Membership was not concealed from the participants or the clinicians; however, membership was concealed from the analyzers. In order to be included in the study, all participants had to have a diagnosis of autism by a physician. The ages of the participants ranged from 3-5 years, with a mean of 4.8. The author included vague information about the participants' cognitive skills by categorizing them into low and high functioning using scores on the *Childhood Autism Rating Scale* or the *Autism Diagnostic Interview Revised*. Expressive and receptive language skills were provided through language age equivalents ranging from 1-4 on the *Preschool Language Scale*, *Peabody Picture Vocabulary Test*, and the *Expressive and Receptive One Word Picture Vocabulary Test*. The author did not include any information regarding gender, mean length of utterance, socioeconomic status, educational level of clients, or educational level of parents. It is unclear if the groups were similar prior to intervention. Throughout the study, the groups maintained 100% of original members.

The time involved in the comparison and target groups were not constant. All participants went through six weeks of either the music or training condition.



The music condition was nine minutes long and was conducted live in a therapy session. The speech condition was five minutes and forty seconds long and was conducted through a video. Both conditions were presented twice a day for six weeks. The dependent variable in this study was the posttest scores of the *VPES*. This scale rated verbal production of the target word and language components of semantics, phonology, pragmatics, and prosody. These variables were analyzed in a subjective measure. The author reported interrater reliability for the *VPES* to be .999.

Results were analyzed using an ANCOVA to determine if there were any significant effects. Both training conditions (speech and music) had a significant effect on participants' verbal production with a p value of  $<.001$ . Music and speech training had a large effect compared to the no training condition. The music condition had a d value of 1.1275 and the speech condition had a d value of 1.141. The author also compared the results of high functioning participants versus low functioning participants. Participants with a high level of functioning performed better than those with low functioning, indicated by a d value of 1.605. There was no statistical difference between participants in the music and speech conditions on the semantics, phonology, pragmatics, and prosody with p values ranging from .709 to .995.

**Critique.** The design of prospective, randomized groups with controls was beneficial to look at a no treatment group to compare the effectiveness of this specific intervention. More information regarding the groups would have been

beneficial. The only information was age range and either low or high functioning autism. With little information about gender, cognitive skills, and language skills it makes the study difficult to apply clinically to specific populations. The author also reported an average language age on some language tests. Language ages do not equate well clinically.

The author provided a sufficient amount of data to replicate the study. Information about the songs sang, the way the stimuli were presented, and targets were displayed clearly in tables and narratively. However, the time involved and the presentation of the comparison and target groups were not consistent. The music condition was nine minutes long and live while the speech condition was a video was five minutes and forty seconds long. Over the six week period this would add up to a major time difference and it would be assumed that better outcomes would be produced in a live treatment session than watching a video. Due to the larger amount of time invested in the music condition it can be viewed as disappointing that the music condition was not significantly better. The pretest and posttest instrument used was the *VPES*, which the author of this study created. We have no information on the validity of this assessment procedure. The measures of the test can be seen as very subjective; whether the instructor believed that a word was produced correctly, or it was said with the correct prosody. To be used in the field of communication disorders, an instrument that is commonly used to assess clients would have been more valid in measuring effectiveness of the treatment. Outcomes reported

showed significant differences and large effects with  $d$  values over 1. However, it did not prove the sung versus the spoken treatment had a more significant effect over the other.

Overall, the final grade for this study was a C+. The treatment groups had a large effect when compared to the no treatment group. It was apparent that the music and speech method used in this study had positive effects on the participants. The design was well developed. However, it is unknown if the form of assessment is an appropriate method to evaluate what is needed in the communication disorders field. The *VPES* example provided in the study had all subjective measurements, which would make it difficult to statistically show effectiveness.

**O'Loughlin (2000).**

**Summary.** The purpose of this study was to investigate if a combination of music and language therapy would increase prelinguistic communication behaviors in children with autism. Observations were made regarding eye contact, looking and pointing at a stimulus, peer engagement, and imitations of talking and singing. The author conducted four separate experiments with no recognizable design to the overall study. The author stated the design was a within subject, repeat measure. There were a total of 44 participants with four separate groups that were not randomly assigned and were all intervention groups. All participants had to have diagnosis of autism by a health care provider. The author provided information about the participants including age

range of 3-20, gender with 37 males and 7 females, and functional level with 18 severely impaired, 12 moderately impaired, and 14 high functioning. The groups did not maintain 80% of their members. Group two retained five of eleven members.

There was no comparison group or control group. The four groups were not designed in a way to compare effectiveness. The outcomes for this research were achieved by graphing the frequency of means for the subjective measured variables of eye contact, looking and pointing at a stimulus, peer engagement, and imitations of talking and singing. Interrater reliability was calculated for two of the four experimental groups. The author reported mean scores for interrater reliability for each observer on each participant in that group. Mean scores were between 0.763 and 0.881.

Evaluation and measurement were completed daily and for a final evaluation using a 5-point *Likert scale* which measured educators' perceptions of the participants' attention. The author included information regarding agreement across observers. This review will only include information regarding speech and language implications. The author reported data on one of the four groups. Group two had a significant difference in the in eye contact ( $p=.023$ ) and looking at the picture stimulus ( $p=.014$ ). No other statistical information regarding participants' performance was provided.

**Critique.** The approach the author took to completing this investigation was very confusing. There was no evident design to these groups. The author

did attempt to assign a design to one of the four groups; however, that approach was still unclear. The author stated the design was a within subject, repeat design. This reviewer could not find an equivalent design of the overall study to assign a grade. No recognizable design to the experiment was a major deficit to this study.

The author did not include enough data about the participants. An age range, gender, and a classification of either low, moderate, or high functioning autism limits generalization due to the lack of participant information. Information regarding language skills and cognitive skills would have been beneficial. The treatment method was described adequately enough to reduplicate.

There were several sections of information regarding interrater reliability; however, it was only reported once. The author placed more emphasis on the agreement across observers for all groups than on the performance of the participants. The dependent variables of eye contact, looking and pointing at a stimulus, peer engagement, and imitations of talking and singing were subjective measurements of performance. It was difficult to appreciate fully values presented when the variables were so subjective because it would be difficult to classify what exactly variables like peer engagement meant. The author reported data in a narrative format with no evidence based practice measurements. The author claimed the music intervention in conjunction with language therapy may help develop attention skills in children with autism, but there were no evidence based measurements behind the claim.

The overall grade for this study was a D-. The author provided a large amount of information to support the reduplication of the study. However, the lack of data for the participants' performance and no evidence based practice measurements given negatively impact the quality of this experiment.

## Chapter IV

### **Discussion**

This thesis presents a systematic review of research articles that evaluated intervention practices utilizing music in speech and language therapy for preschool children and children with autism. Overall, the studies revealed a general consensus that music in therapy does improve language skills in preschool children and children with autism. However, the data should be viewed with caution due to poor study designs, methods of obtaining outcomes, and the means used to interpret data.

#### **Preschool Based Studies**

The search yielded a total of six studies designed to treat preschool aged children using music in speech and language therapy. The results of these studies provided general positive outcomes to improving the speech and language skills of preschoolers using music therapy. Table 2 provides a brief summary of the reviewer's overall grade assigned to the study based on the general quality of the study. Of the six studies, the review and critique yielded four studies with Cs as an overall grade and two studies with Ds. The general low grades were due to the lack of EBP measures in data analysis and interpretation. The study that received the highest grade was Kouri & Winn (2006) because of the use of EBP metrics and significant differences in some of the outcomes measured. The study that received the lowest grade was Seeman (2008)

Table 2

*Reviewer's overall grading for preschool based studies*

Author	Findings	Grade
Aldridge, Gustroff, & Neugebauer (1995)	The outcomes for hearing and speech were p=0.004, outcomes for personal-social were p=0.044, and outcomes for practical reasoning were p=0.188	D
Gross, Linden, & Ostermann (2010)	No effect for generation of morphological rules and memory for word sequences, small effect for phonological memory for nonwords and understanding sentences, moderate effect for memory for sentences	C
Hoskins (1988)	High ability participants were significantly better on the spoken <i>PPVT</i> ( $p<.05$ ) from the low ability participants. <i>PPVT</i> melodic high ability group was significantly different ( $p<.05$ ). Significant improvement for all participants ( $p<.05$ ) with no significant group by trials interaction. Pretest and posttest which showed a significant improvement ( $p<.05$ ). <i>EOWPVT</i> results high ability group was significantly different ( $p<.05$ )	C-



Kouri & Winn (2006)	Large effect on unsolicited production and large effect correlation between condition and number of sessions	C+
Ross (1997)	Highly effective for production of /m/, ineffective for production of /p/, and questionable effectiveness for /b/	C-
Seeman (2008)	21.18% age equivalent increase on the <i>PPVT</i> and an average of 34.67% increase of age equivalent scores on the <i>TROLL</i>	D-

---

*Note.* *PPVT* = Peabody Picture Vocabulary Test. *EOWPVT* = Expressive One Word Picture Vocabulary Test. *TROLL* = Teacher Rating of Oral Language and Literacy.

because of the use of age equivalents to report outcomes and no form of EBP metrics was used.

The studies reviewed concluded that using music therapy with this population had positive effects on the following conditions: expressive speech including production of the /m/ and /b/ phonemes, expressive vocabulary, phonological memory for nonwords, understanding sentences, memory for sentences, comprehension of novel items, production of novel items, expressive language skills, and personal and social behaviors. The authors concluded that music therapy did not appear to improve the production of the /p/ phoneme, generation of morphological rules, memory for word sequences, and practical reasoning.

### **Autism Based Studies**

The search yielded a total of three studies designed to treat children with autism using music in speech and language therapy. The results of these studies had mixed overall effectiveness. Table 3 provides a brief summary of the reviewer's overall grade assigned to the study based on the general quality of the study. Of the three studies, the review and critique yielded one study with C as an overall grade and two studies with Ds. The general low grades were due to the measure of outcomes chosen to show effectiveness and the lack of EBP measures in data collection and interpretation. The study that received the highest grade was Lim (2010) because of the type of evidence used, the use of a comparison group, and the general positive effects. The study that

Table 3

*Reviewer's overall grading for autism based studies*

Author	Findings	Grade
Edgerton (1994)	Last session scores on the <i>CRASS</i> were greater than first session scores (.01). Significant differences were found at the .05 level between first and last session scores for communicative intent. <i>The Behavior Change Survey</i> significant both communication and social/emotional categories received a mean of 4.5 (4=no change 5=slight change)	D+
Lim (2010)	Training conditions had a significant effect on verbal production ( $p < .001$ ) compared to no treatment ( $d = 1.1275$ ). High functioning performed better than low functioning participants ( $d = 1.605$ )	C+
O'Loughlin (2000)	Significant difference ( $p = .025$ ) for eye contact and looking at the picture stimulus ( $p = .014$ )	D-

*Note.* *CRASS* = The Checklist of Communicative Responses/Acts Score Sheet.

received the lowest grade was O'Loughlin (2000) because of the unclear design approach, lack of participant information, and the lack of EBP metrics.

The results of these studies reviewed concluded that using music therapy with this population had positive effects on the following conditions:

communicative intent, speech production (especially in high-functioning participants), eye contact, and looking at picture stimulus. The authors concluded that music therapy did not appear to improve phonology, pragmatics, prosody, peer engagement, imitation of talking, and social and emotional behaviors.

### **Clinical Implications**

SLPs who provide intervention to students using music therapy may find the results somewhat contradictory and concerning. The fact that only nine studies met the criteria for inclusion in this review signifies how little research has been conducted on this rapidly growing form of intervention. Clinicians need to ensure that clients receive services that are known to be based on reliable and valid research. The lack of evidence on utilizing music in speech and language therapy limits the ability to use EBP in the clinical setting. In order to provide high-quality services to clients it is vital that SLPs use forms of intervention that have strong efficacy, effectiveness, and efficiency. It is recommended that if SLPs choose to use music in therapy, it should be done in conjunction with EBP methods of treatment that provide evidence of effectiveness until further research is conducted on this topic or that they regularly and closely monitor clients'

progress and/or validate the program's effectiveness with their clients using single subject experimental designs.

Overall, the nine studies that were reviewed and critiqued showed a general positive change in speech and language outcomes. Tables 4 and 5 summarize overall findings of the most successful outcomes in the reviewed articles to help illustrate what behaviors using music in speech and language therapy has shown a positive effect. Based on the information provided in tables 4 and 5, the outcomes most likely to be responsive to utilizing music in speech and language therapy are unsolicited production of novel items, phonological memory for nonwords, understanding sentences, memory for sentences, scores on the *PPVT*, scores on the *EOWPVT*, production of /m/ and /b/, hearing, personal-social behaviors, age equivalent scores on the *PPVT* and *TROLL*, verbal production, scores on the *CRASS* and *The Behavior Change Survey*, eye contact, and looking at picture stimulus.

The purpose of this thesis was to determine if there were high-quality studies that documented the effectiveness of using music in speech and language therapy with preschool aged children and children with autism. After reviewing the nine studies in this thesis, it can be concluded that high-quality evidence does not exist to show the effectiveness of using music in speech and language therapy with the targeted populations. The studies did show overall improvement on the measured outcomes; however, the overall quality of the designs was determined low.

Table 4

*Clinical implications of preschool based results*

Authors	Significant positive outcomes	Reviewer's grade
Kouri & Winn (2006)	Unsolicited production	C+
Gross, Linden, & Ostermann (2010)	Phonological memory for nonwords, understanding sentences, and memory for sentences	C
Hoskins (1988)	Scores on <i>PPVT</i> (spoken and melodic) and <i>EOWPVT</i>	C-
Ross (1997)	/m/ and /b/	C-
Aldridge, Gustroff, & Neugebauer (1995)	Hearing and speech and personal- social behaviors	D
Seeman (2008)	Age equivalent scores on the <i>PPVT</i> and <i>TROLL</i>	D-

*Note.* PPVT = Peabody Picture Vocabulary Test. EOWPVT = Expressive One Word Picture Vocabulary Test. TROLL = Teacher Rating of Oral Language and Literacy.

Table 5

*Clinical implications for autism based results*

Author	Significant positive outcomes	Reviewer's grade
Lim (2010)	Verbal production	C+
Edgerton (1994)	<i>CRASS</i> and <i>The Behavior Change Survey</i> scores	D+
O'Loughlin (2000)	Eye contact and looking at picture stimulus	D-

*Note.* CRASS = The Checklist of Communicative Responses/Acts Score Sheet.

## **Future Research**

Increased consultation and collaboration between SLPs and music therapists may increase the applicability and effectiveness of using music in speech and language therapy. Combining the two professions could possibly create specified techniques that incorporate both music and speech-language therapy to improve client goals.

Many results from the initial search were descriptions of curricula using music therapy. These descriptions of curricula did not include any evidence to support the methods or to show effectiveness. With the growing interest in utilizing music in speech and language therapy, it is vital that researchers conduct studies that provide EBP metrics to allow clinicians to incorporate music within speech and language therapy in an ethical manner. It is noted that not all of the authors of the studies reviewed in this thesis were SLPs. Larger sample sizes need to be included within future studies to facilitate better generalization of the experimental techniques. In order for SLPs to utilize music within speech and language therapy with children, future research needs to be conducted by professionals within the field. Research by SLPs might result in the selection of outcomes that are more relevant to the field and, therefore, support the use of EBP within the clinical SLP setting.

Future studies need to include characteristics that many of the studies in this thesis lacked in order to satisfy the needs to increase EBP practice in the clinical setting to use music in speech and language therapy. The following list



includes, but is not limited to, suggestions of characteristics to ensure studies will contain more positive evidence to using this method in the clinical SLP setting.

- Studies with strong levels of support (i.e., randomized clinical trials)
- Random assignment to group membership
- Group membership that is concealed from participants, clinicians, and analyzers
- Adequate description of group members (i.e., age, gender, cognitive skills, expressive language skills, receptive language skills, mean length of utterance, socioeconomic status, educational level, and educational level parents)
- Communication problems described adequately (i.e., disorder type, functional level, the use of standardized tests to describe disorder)
- Control group(s)
- Appropriate outcomes relevant to the field of speech-language pathology
- Reliability measures
- Use of EBP measures to illustrate clinical effect
- Adequate description of intervention used in the study to generalize in the clinical setting

## References

- Aldridge, D., Gustroff, D., & Neugebauer, L. (1995). A pilot study of music therapy in the treatment of children with developmental delay. *Complementary Therapies in Medicine*, 3(4), 197-205.
- American Speech-Language-Hearing Association. Introduction to evidence-based practice. Available from [www.asha.org/Members/ebp/intro](http://www.asha.org/Members/ebp/intro)
- American Speech-Language-Hearing Association. (2005). *Evidence-based practice in communication disorders* [Position Statement]. Available from [www.asha.org/policy](http://www.asha.org/policy). doi:10.1044/policy.PS2005-00221
- American Speech-Language-Hearing Association. (1993). *Definitions of communication disorders and variations* [Relevant Paper]. Available from [www.asha.org/policy](http://www.asha.org/policy). (doi:10.1044/policy.RP1993-00208)
- Arntson, R. (2006). *Music, the underutilized tool in speech and language therapy*. Presented at the American Speech-Language-Hearing Association convention, Miami, FL. Retrieved from [http://www.asha.org/Events/convention/handouts/2006/1495\\_Arntson\\_Rachel.htm](http://www.asha.org/Events/convention/handouts/2006/1495_Arntson_Rachel.htm)
- Arntson, R. (2009). *Music and autism: Connections in a disconnected world*. Presented at the American Speech-Language-Hearing Association convention, New Orleans, LA. Retrieved from [http://www.asha.org/Events/convention/handouts/2009/1265\\_Arntson\\_Rachel.htm](http://www.asha.org/Events/convention/handouts/2009/1265_Arntson_Rachel.htm)
- Boucher, M. (2008). *Music therapy and enhancing speech therapy sessions with music*. Presented at the American Speech-Language-Hearing Association

convention, Chicago, IL. Retrieved from [http://www.asha.org/Events/convention/handouts/2008/2162\\_Boucher\\_Michele.htm](http://www.asha.org/Events/convention/handouts/2008/2162_Boucher_Michele.htm)

Edgerton, C. L. (1994). The effect of improvisational music therapy on the communicative behaviors of autistic children. *Journal of Music Therapy*, 21(1), 31-62.

Gross, W., Linden, U., & Ostermann, T. (2010). Effects of music therapy in the treatment of children with delayed speech development - Results of a pilot study. *BMC Complementary and Alternative Medicine*, Jan 01, 39. Retrieved from <http://www.biomedcentral.com/1472-6882/10/396882/10/39>.

Hoskins, C. (1988). Use of music to increase verbal response and improve expressive language abilities of preschool language delayed children. *Journal of Music Therapy*, 25 (2), 73-84.

Kouri, T. A., & Winn, J. (2006). Lexical learning in sung and spoken story script contexts. *Child Language Teaching and Therapy*, 22, 293-313.

Lim, H. A. (2010). Effect of developmental speech and language training through music on speech production in children with autism spectrum disorders. *Journal of Music Therapy*, 47(1), 2-26.

McCarthy, J. & Geist, K. (2007). *A survey of music therapists' collaboration with speech- language pathologists*. Presented at the American Speech-Language-Hearing Association convention, Huntersville, NC. Retrieved

from [http://www.asha.org/Events/convention/handouts/2007/1710\\_McCarthy\\_John.htm](http://www.asha.org/Events/convention/handouts/2007/1710_McCarthy_John.htm)

McCauley, R. & Hargrove, P. (2004). A clinician's introduction to systematic review in communication disorders: The course review paper with muscle. *Contemporary Issues in Communication Science and Disorders*, 31, 173-181.

Minnesota Department of Education. (2009). *Early childhood 2009 annual report: Minnesota early childhood advisory council*. Retrieved from <http://education.state.mn.us/mdeprod/groups/EarlyLearning/documents/Announcement/014622.pdf>

Nordoff Robbins. (2011). Nordoff Robbins – Music transforming lives. Retrieved from <http://www.nordoff-robbins.org.uk/>

O'Loughlin, R. A. (2000). *Facilitating prelinguistic communication skills of attention by integrating a music stimulus within typical language intervention with autistic children*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 9965033).

Paul, R. (2007). *Language disorders from infancy through adolescence*. St. Louis, MO: Mosby Elsevier.

Ross, S. Y. (1997). *Effects of singing on speech patterns of children with expressive language delays*. (Master's thesis). Retrieved from ComDisDome. MAI 36/01,. (cs-6064; 1386398).

- Sackett, D. L. (2000). *Evidence based medicine: How to practice and teach EBM* (2<sup>nd</sup>ed.). Edinburgh: Churchill Livingstone.
- Schiavetti, N. & Metz, D. E. (2006). *Evaluating research in communicative disorders*. Boston, MA: Pearson Education, Inc.
- Seeman, E. (2008). *Implementation of music activities to increase language skills in the at-risk early childhood population*. (Master's thesis). Retrieved from ERIC. (ED503341).
- Zoller, M. B. (1991). Use of music activities in speech-language therapy. *Language, Speech, and Hearing Services in Schools* 21, 272-276.

## Appendix A

## EBP THERAPY ANALYSIS

SOURCE:

REVIEWER(S):

DATE:

ASSIGNED GRADE for QUALITY:

QUESTIONS	Comments
1. <i>What type of evidence was identified?</i>	
1a. What was the type of evidence? (circle type) <ul style="list-style-type: none"> <li>• Randomized Clinical Trial (RCT)</li> <li>• Systematic Review (SR)</li> <li>• Prospective, Randomized Group Design with Controls (PRG)</li> <li>• Single Subject Experimental Design with Specific Client (SSED)</li> <li>• Prospective, Nonrandomized Group Design with Controls? (PNG)</li> <li>• Evidence-Based Clinical Practice Guidelines (EBCPG)</li> <li>• Retrospective, Nonrandomized Group Design with Controls? (RNG)</li> <li>• Prospective, Single Group with Pre- and Post-Testing (PSG)</li> <li>• Traditional Clinical Practice Guideline (TCPG)</li> <li>• Case Series (CSe)</li> <li>• Case Studies (CSt)</li> <li>• Narrative Literature Review (NLR)</li> <li>• Descriptive Research (DR)</li> <li>• Essential Research (ER)</li> <li>• Expert Opinion (EO)</li> </ul>	
1b. What was the level of support associated with the type of evidence?	Level = _____
2. <i>How was group membership determined?</i>	
2a. Were participants randomly	yes _____ no

QUESTIONS		Comments
assigned to groups?	_____ unclear _____	
2b. If subjects were not randomly assigned to groups, were members of groups carefully matched?	yes _____ no _____ unclear _____	
2c. If the answer to 2a and 2b is 'no' or 'unclear,' describe assignment strategy:		
3. <i>Was group membership concealed?</i>		
a. from participants?	yes _____ no _____ unclear _____	
b. from clinicians?	yes _____ no _____ unclear _____	
c. from analyzers?	yes _____ no _____ unclear _____	
4. <i>Were the groups adequately described?</i>		Yes ___ No ___ Unclear _____
4a. How many subjects were involved in the study? total # of participants # of groups	_____ _____ _____	
4b. Were the following variables actively controlled (i.e., inclusion/exclusion criteria) or determined to be similar? (check variables that are actively controlled or similar)		
	control	similar
i. age?		
ii. gender?		
iii. cognitive skills		

iv. expressive language?			
v. receptive language?			
vi. MLU?			
vii. SES?			
viii. educational level of clients?			
ix. educational level of parents?			
x. age at referral			
Other (list):			
4c. Were the groups similar before intervention began?			
	yes _____ no _____ unclear _____		
4d. Were the communication problems adequately described?			
i. disorder type?	yes _____ no _____ unclear _____	List:	
ii. functional level?	yes _____ no _____ unclear _____	List:	
iii. other (list)	yes _____ no _____ unclear _____		
5. Was membership in groups maintained throughout the study?			
a. Did each of the groups maintain at least 80% of their original members?	yes _____ no _____ unclear _____		
b. Were data from outliers removed from the study?	yes _____ no _____ unclear _____		
6. Were the groups controlled acceptably?			
a. Was there a no intervention	yes _____ no _____ unclear _____		



group? (NI)				
b. Was there a foil intervention group? (FI)	yes _____ no _____ unclear _____			
c. Was there a comparison group? (CI)	yes _____ no _____ unclear _____			
d. Was the time involved in the foil/ comparison and the target groups constant?	yes _____ no _____ unclear _____			
<i>7. Were the outcomes measure appropriate and meaningful?</i>				
7a. List outcome (dependent variable):				
7b. Is the outcome measure subjective?	Yes	No	Unclear	
7c. Is the outcome measure objective?	Yes	No	Unclear	
<i>8. Were reliability measures provided?</i>				
a. Interobserver for analyzers? (Inter)	yes _____ no _____ unclear _____			If yes, list:
b. Intraobserver for analyzers? (Intra)	yes _____ no _____ unclear _____			If yes, list:
c. Treatment fidelity for clinicians? (Fidel)	yes _____ no _____ unclear _____			If yes, list:
<i>9. What were the results of the statistical (inferential) testing?</i>				
9a. List the order of improvement on the outcome measure	1st	2nd	3rd	4 <sup>th</sup> or more

from most to least improvement:				
9b. Was there a significant difference in outcome measures following treatment?	Yes	No	Unclear/ Variable	
9c. What was the p value?	t-test	ANOVA	other	
9d. Was confidence interval (CI) provided?	Yes	No	Unclear/ Variable	
9e. What is the percentage associated with the confidence interval (CI)?	98%	95%	90%	List CI if not one of 3 provided:
9f. List CI (range) under appropriate percentage:				
<i>10. What is the clinical effect? (i.e., EBP measures; check measure reporting)</i>				
___ Standardized Mean Difference ___ Effect Size ___ Correlation ___ Number Needed to Treat ___ ETA ___ $r^2$ ___ other	<u>Score</u>	<u>CI</u>	<u>Interpretation</u>	

ASSIGNED GRADE FOR QUALITY OF EXTERNAL EVIDENCE: \_\_\_\_\_

## Appendix B

## Types of evidence, description, and assigned grade

Type of Evidence	Description	Grade
Randomized Clinical Trial (RCT)	Prospective, randomized group design with a control group and sufficient subjects to allow for a small confidence interval.	A+
Systematic Review	A comprehensive, critical analysis of the literature pertaining to a specific treatment strategy. Rational for inclusion/exclusion of research is provided. Studies limited to RCTs and Prospective Randomized, Groups with Controls for the treatment studies and Prospective, Nonrandomized groups with controls.	A+
Prospective, Randomized Group with Controls	Subjects are randomly assigned to experimental and control groups prior to treatment of the experimental groups. Outcomes before and after treatment are monitored for both groups.	A
Single Subject Experimental Design with Specific Client	The SLP employs a single subject experimental design to determine the effectiveness of a treatment procedure with his/her client.	B+
Systematic Reviews	A comprehensive, critical analysis if the literature pertaining to a specific treatment strategy. Rational for inclusion/exclusion of research is provided. Studies reviewed can include a variety of designs.	B
Prospective, Single Group Design with Pre- and Post-testing	Subjects are assigned to an experimental group prior to treatment. Outcomes before and after treatment are monitored for the group.	B-
Retrospective, Nonrandomized Group Design with Controls	Group membership is assigned after the outcome is known. Attempt is made to determine variable(s) associated with the desired outcome	C+

## Appendix C

Utilizing music in speech and language therapy focused on preschool aged children

Citation	Participants ; disorder	Research design; dependent variables	Major findings	Review Grade
Aldridge, D., Gustroff, D., & Neugebauer, L. (1995). A pilot study of music therapy in the treatment of children with developmental delay. <i>Complementary Therapies in Medicine</i> , 3(4), 197-205.	8 participants ages 4.0-6.5; Developmental delay	Prospective, randomized group design with controls; <i>Griffiths</i> test subscale scores	The outcomes for hearing and speech were $p=0.004$ , outcomes for personal-social were $p=0.044$ , and outcomes for practical reasoning were $p=0.188$	D
Gross, W., Linden, U., & Ostermann, T. (2010). Effects of music therapy in the treatment of children with delayed speech development - Results of a pilot study. <i>BMC Complementary and Alternative Medicine</i> , Jan 01, 39. Retrieved	18 participants ages 6.5-6; Speech delay	Prospective, single group design with pre- and posttest; Phonological memory for nonwords, memory for sentences, generation of morphological rules, and memory for word sequences	No effect for generation of morphological rules and memory for word sequences, small effect for phonological memory for nonwords and understanding sentences, moderate effect for memory for sentences	C

from  
<http://www.biomedcentral.com/1472-6882/10/396882/10/39>.

<p>Hoskins, C. (1988). Use of music to increase verbal response and improve expressive language abilities of preschool language delayed children. <i>Journal of Music Therapy</i>, 25 (2), 73-84.</p>	<p>16 participants ages 2.0-5.0 years; Developmental delay</p>	<p>Prospective, single group with pre- and posttesting; <i>PPVT</i> and <i>EOWPVT</i> scores</p>	<p>High ability subjects were significantly different on the spoken <i>PPVT</i> (<math>p &lt; .05</math>) from the low ability subjects. <i>PPVT</i> melodic high ability group was significantly different (<math>p &lt; .05</math>). Significant improvement for all subjects (<math>p &lt; .05</math>) with no significant group by trials interaction. Pretest and posttest which showed a significant improvement (<math>p &lt; .05</math>). <i>EOWPVT</i> results high ability group was significantly different (<math>p &lt; .05</math>).</p>	<p>C-</p>
<p>Kouri, T. A., &amp; Winn, J. (2006). Lexical learning in sung and spoken story script contexts. <i>Child Language Teaching and</i></p>	<p>16 participants ages 3.6-5.1 years; Developmental delay</p>	<p>Prospective, single group with pre- and post testing; Comprehension, production, generalization, and unsolicited</p>	<p>Large effect on unsolicited production and large effect correlation between condition and number of session</p>	<p>C+</p>

<i>Therapy</i> , 22, 293-313.			production of novel items.	
Ross, S. Y. (1997). <i>Effects of singing on speech patterns of children with expressive language delays</i> . (Master's thesis). Retrieved from ComDisDome. MAI 36/01,. (cs-6064; 1386398)	3 participants ages 3-5 years; Severe expressive speech impairment	Single subject experiment with multiple baseline; Production of /p/, /b/, and /m/	Highly effective for production of /m/, ineffective for production of /p/, and questionable effectiveness for /b/	C-
Seeman, E. (2008). <i>Implementation of music activities to increase language skills in the at-risk early childhood population</i> . (Master's thesis). Retrieved from ERIC. (ED503341).	9 participants ages 3.6-4.1; At-risk or special needs	Prospective, single group with pre- and post-testing; Age equivalent scores on the PPVT and TROLL	21.18% age equivalent increase on the PPVT and an average of 34.67% increase of age equivalent scores on the TROLL	D-

---

*Note.* PPVT = Peabody Picture Vocabulary Test. TROLL = Teacher Rating of Oral Language and Literacy. EOWPVT = Expressive One-Word Picture Vocabulary Test.

## Appendix D

Utilizing music in speech and language therapy focused on children with autism

Citation	Participants; functioning range	Research design; dependent variables	Major findings	Review Grade
Edgerton, C. L. (1994). The effect of improvisational music therapy on the communicative behaviors of autistic children. <i>Journal of Music Therapy, 21</i> (1), 31-62.	11 participants ages 6-9 years; Severely impaired-mildly impaired	Single subject experimental design with specific clients; <i>CRASS &amp; The Behavior Change Survey</i>	Last session scores on the <i>CRASS</i> were greater than first session scores (.01). Significant differences were found at the .05 level between first and last session scores for communicative intent. <i>The Behavior Change Survey</i> significant both communication and social/emotional categories received a mean of 4.5 (4=no change 5=slight change)	D+

<p>Lim, H. A. (2010). Effect of developmental speech and language training through music on speech production in children with autism spectrum disorders. <i>Journal of Music Therapy, 47</i>(1), 2-26.</p>	<p>50 participants ages 3-5 years; Low-high functioning</p>	<p>Prospective, randomized group design with controls; VPES (designed by author)</p>	<p>Training conditions had a significant effect on verbal production (<math>p &lt; .001</math>) compared to no treatment (<math>d = 1.1275</math>). High functioning performed better than low functioning participants (<math>d = 1.605</math>)</p>	<p>C+</p>
<p>O'Loughlin, R. A. (2000). <i>Facilitating prelinguistic communication skills of attention by integrating a music stimulus within typical language intervention with autistic children.</i> (Doctoral dissertation). Retrieved from Dissertations &amp; Theses: Full Text. (Publication No. AAT 9965033).</p>	<p>44 participants ages 3-10 years; Severely impaired-high functioning</p>	<p>No clear design; Likert's scale (eye contact, looking and pointing at stimulus, peer engagement, imitations of talking/singing)</p>	<p>Significant difference (<math>p = .025</math>) for eye contact and looking at the picture stimulus (<math>p = .014</math>)</p>	<p>D-</p>

---

Note. CRASS = The Checklist of Communicative Responses/Acts Score Sheet.