



# Musical Mnemonics Enhance Verbal Memory in Typically Developing Children

David Knott<sup>1\*</sup> and Michael H. Thaut<sup>2</sup>

<sup>1</sup> Child Life Department, Seattle Children's Hospital, Seattle, WA, United States, <sup>2</sup> Faculty of Music, Music and Health Science Research Center, University of Toronto, Toronto, ON, Canada

## OPEN ACCESS

### Edited by:

Meryem Yilmaz Soylu,  
University of Nebraska-Lincoln,  
United States

### Reviewed by:

Angela Jocelyn Fawcett,  
Swansea University, United Kingdom  
Benjamin Becker,  
University of Electronic Science and  
Technology of China, China

### \*Correspondence:

David Knott  
david.knott@seattlechildrens.org

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

Received: 12 December 2017

Accepted: 25 April 2018

Published: 15 May 2018

### Citation:

Knott D and Thaut MH (2018) Musical  
Mnemonics Enhance Verbal Memory  
in Typically Developing Children.  
Front. Educ. 3:31.  
doi: 10.3389/feduc.2018.00031

The purpose of this study was to compare the effects of musical mnemonics vs. spoken word in training verbal memory in children. A randomized control trial of typically-developing 9–11 year old children was conducted using the Rey Auditory Verbal Learning Test (RAVLT), a test measuring a participant's ability to recall a list of 15 words over multiple exposures. Members of the group who listened to words sung to them recalled an average of 20% more words after listening to and recalling an interference list than members of the control group who listened to the same words spoken. This difference persisted, though slightly smaller (17%) when participants recalled words after a 15-min waiting period. Additionally, group participants who listened to words sung demonstrated a higher incidence of words recalled in correct serial order. Key findings were all statistically significant at the  $P < 0.05$  level. Enhanced serial order recall points to the musical pitch/rhythm structure enhancing sequence memory as a potential mnemonic mechanism. No significant differences were found in serial position effects between groups. The findings suggest that musical mnemonic training may be more effective than rehearsal with spoken words in verbal memory learning tasks in 9–11 year olds.

**Keywords:** music, mnemonics, verbal memory, RAVLT, children, learning, songwriting, music therapy

## INTRODUCTION

Memory is a critical component of cognitive functioning and deficits to verbal memory are a feature of many neurological disorders and injuries, including traumatic brain injury, brain tumor, epilepsy, stroke, developmental disability, autism, Attention Deficit Hyperactivity Disorder (ADHD) and Down syndrome (Dehn, 2010). Furthermore, there is some indication that children with poor literacy may also have deficits in memory function (Al Otaiba and Fuchs, 2006). Musical mnemonics, or using rhythmic-melodic templates for rehearsing verbal information, provides a uniquely engaging approach to verbal memory tasks. Despite the widespread use of this technique for teaching information such as the alphabet and for students in higher education to memorize complex material (Cirigliano, 2013), there is little research on the use of musical mnemonics. Increased understanding of the effects of this learning strategy may lead to improved academic performance for typically developing children, those who struggle with traditional approaches to learning, and individuals with deficits in memory due to injury or illness. Musical mnemonics may also provide clinical populations a simple strategy for memorizing important safety information such as a home phone number or address.

As meaning is often derived from the serial presentation of information, an important consideration in memory research is serial location and order. In trials of verbal memory, primacy refers to words initially presented in a learning list and primacy was found to be affected by memory dysfunction (Lezak, 1983). In normative trials with typically-developing 9–10 year-olds, serial position effects were noted for both primacy and recency (words presented at the end of the list; Forrester and Geffen, 1991). Given these known patterns of learning with both memory disordered and typical populations, it is also important to consider the effects of serial order in learning using musical mnemonics. Studies of individuals without hearing revealed decreased prefrontal cortical development as well as poorer performance on tests of sequencing than age matched controls with normal hearing (Conway et al., 2009). These findings suggest a privileged role for auditory information in the learning process. Music, as a system for delivering highly organizing auditory information, may have unique effects upon operations involving the phonological loop (Baddeley and Hitch, 1974). Of special importance in this respect for music as a mnemonic device have been early experiments by Miller (1956) which revealed that working memory processes allow us to group information together into meaningful chunks, often based on relationships with the individual's long term memories. As music is an entirely time ordered auditory language that utilizes a hierarchy of notes, phrases, motives and rhythms to organize and group information (Deutsch, 1982), it may provide additional structure in working memory tasks akin to Miller's "chunking" observations Miller (1956).

Musical training has been shown to change brain structures and summaries of imaging studies conducted over the past 20 years corral evidence suggesting that regular study of music influences plastic changes in brain functioning (Habib and Besson, 2009; Wan and Schlaug, 2010; Moreno et al., 2011). Correlational studies have found links between memory performance and musical training. Chan et al. (1998) found that adults who had musical training before the age of 12 had significantly better verbal memory performance than age-matched peers without early musical training. A later study conducted by Ho et al. (2003) found similar results with a child population. George and Coch (2011) compared event-related potential (ERP) data and performance on a standardized test of working memory of musicians and non-musicians and found both better performance measures as well as ERP data suggesting more efficient working memory performance in musicians. An experimental study of 4–6 year old children engaged in 1 year of Suzuki-based violin instruction performed better and improved more significantly than their age matched peers in digit span memory tests (Fujioka et al., 2006). Taken together, these correlational and experimental data support theories that musical training has a transfer effect upon a participant's working memory ability. But targeted training of verbal memory using music can have more immediate effects on learning and verbal memory performance.

An early study by Gfeller (1983) with children with learning disabilities found improved learning of verbal information when presented sung. Further study of training with musical templates was conducted by Wolfe and Hom (1993). In their study with 5 year olds, telephone numbers were presented either

spoken or sung using a familiar melody. Study participants in the sung condition required significantly fewer trials to learn the target telephone number. Wallace (1994) conducted experiments comparing spoken text vs. sung text and found greater recall when text was presented using the rhythm and melody. Interestingly, when a line of text was sung only once, or if different melodies were used with each repetition of text, then spoken text enabled/led to improved recall. To understand the effects of musical templates on children with learning disabilities, Claussen and Thaut (1997) presented multiplication tables in spoken or sung forms. As in earlier trials, participants receiving the sung condition outperformed their peers who heard the spoken presentation of multiplication tables.

To better understand potential mechanisms of verbal learning training using music, Peterson and Thaut (2007) used electroencephalograph (EEG) measures to compare a sung version of the RAVLT with the conventional spoken version in a randomized trial with 18–26 year olds. The study authors found an increased oscillatory synchronization both within and between left and right prefrontal cortical areas in the group training with sung presentations of words. Study authors arrived at their dependent variable of learning related change in coherence (LRCC) (p. 218) by measuring synchronous firing of proximal pairs of scalp electrodes in one hemisphere with synchronous firing in the opposite hemisphere when a new word was recalled, suggesting a key physiological difference in brain activity during verbal learning using musical templates vs. spoken templates.

Another study with individuals with Multiple Sclerosis utilizing a similar test procedure while also examining EEG activity found word order memory as predictive of higher overall word recall performance (Thaut et al., 2014). This study also found increased bilateral synchronization of the prefrontal cortices in participants who heard the word lists sung compared to those who heard the lists spoken. Termed by the authors as "learning-related synchronization," this observation also corresponded with improved word order and overall word recall performance, again suggesting temporal ordering of information as potential mechanism for musical templates as mnemonic devices.

Building on previous studies of musical mnemonics, the present study investigated the effect of musical mnemonics in typically developing children using assessments and procedures previously primarily researched with adults and clinical populations by utilizing a musical translation of the RAVLT which has been widely used for quantifying verbal learning and memory across a range of ages (Schmidt, 1996). This study with typically developing children demonstrates the immediate effects of verbal memory training with rhythmic and melodic templates and provides a baseline for further investigation with special populations.

## MATERIALS AND METHODS

### Participants

A sample ( $n = 32$ ) of typically developing children ages 9–11 was recruited from school districts in a major metropolitan area and randomly assigned to listen to either the sung or

spoken condition. Inclusion criteria required participants to be between 9 and 11 years old, or between 108 and 143 months. All parents responded to the call for participants via email or phone and were directed to consider inclusion and exclusion criteria to determine if their child was “typically developing” with normal hearing and without any known developmental or sensory processing disorders. Sample size was estimated using data from Forrester and Geffen’s (1991) large scale study of typically developing Australian children and Thaut et al. (2014) mean change between treatment and control groups of adults with MS. Utilizing these data to inform a power analysis, a similar percentage change in retention score (12.3% or 1.8 words) would require a total  $n$  of 26, or 13 participants per group with an alpha of 0.05 to yield power of 80%. While a total of 33 participants were recruited and participated in the study, one was disqualified from inclusion due to the family reported diagnosis of ADHD. Of the 32 participants included in the analysis, the test and control conditions each contained 16 participants. Test and control condition groups were evenly distributed with 5 girls and 11 boys present in each group and distribution of age was found to be the same across both groups. A block randomization protocol was utilized to assign participants to either the control or music condition. Allocation according to this randomization schedule occurred at the time of participation. Chronology was demonstrated through the codes as each included the date (and a letter if there were multiple participants on a single day) and whether they received the music (Y) or control (X) condition. The experimenters were blinded to assigned conditions.

## Research Design

The study was designed as a randomized and controlled group comparison of the immediate effects of rhythmic and melodic presentation on verbal recall, recognition, serial position and word order. The study was approved by the appropriate institutional review boards. The RAVLT served as the instrument to compare memory performance between sung or spoken

conditions. Novel melodies created by the researcher served as the music used in the test condition.

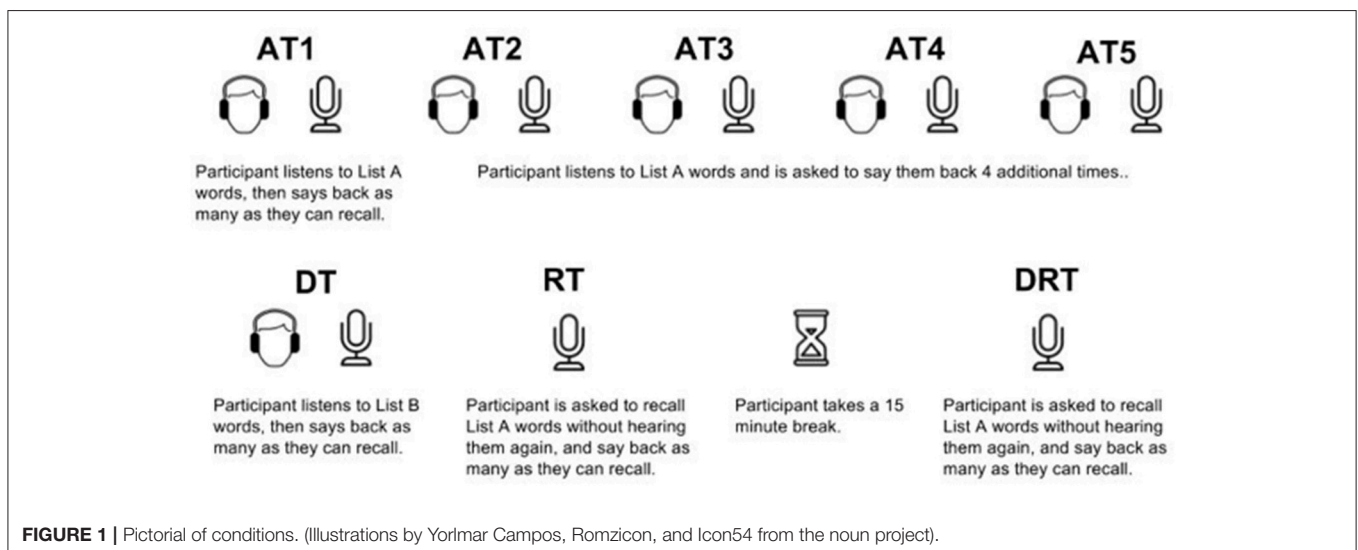
## Measures

Procedures for this study were based on prior research utilizing the Rey Auditory Verbal Learning Test (RAVLT). In keeping with typical administration of the RAVLT, the words were presented one per second and the participant was asked to recall as many words as possible after each presentation. An initial series of five presentations (AT1-5) was followed by presentation of a second list of words that served as a distractor list (DT). After attempting to recall the distraction list words, the participants were again prompted to recall the words from the original list (RT, or retention trial). Participants were then provided a 15-min time break. After the 15-min wash-out period, participants put on headphones and again were prompted by the narrator to recall the original list words (DRT, or delayed recall trial) (see **Figure 1**). After this final retention phase, participants were presented with the Recognition test. The Recognition test was a written form used in prior studies of the RAVLT (Schmidt, 1996), a single page containing 50 words: Words from both lists A and B as well as 20 other words that were not presented on either list.

To evaluate the effect of musical presentation on verbal memory, the Retention trial (RT) was considered the primary outcome. Other data points served as secondary outcomes, to provide insight into possible mechanisms in the musical presentation on various aspects of the learning process.

## Retention and Recall

Secondary dependent variables for this study included summary and composite scores from the first five acquisition trials, delayed recall (DRT following a 15 min wash out interval) and recognition. Additional dependent variables included serial word order and serial position scores. Interference on learning was recorded by counting repetitions (all instances of a word being repeated during recall), intrusions



(words that are recalled by a participant but that do not appear on the target list), and contaminations (words incorrectly recalled as belonging to the opposing list). Two independent inter-raters used the test's recall scoring form to transcribe and score the recording of each participant's performance.

## Acquisition and Learning

*Total learning* (summary of AT 1-5) is a commonly reported measure of RAVLT performance (Forrester and Geffen, 1991; Schmidt, 1996) and provides a summary of participant initial learning before presentation of interference from a distractor list or time delay. In addition to total learning, researchers have developed composite scores in an attempt to isolate the performance of repeated practice, e.g., learning, by removing the influence of the initial response to the words.

In order to better isolate the learning process from the initial recall after the first learning trial, Ivnik et al. (1992) suggested a *corrected total learning* score that eliminates the participants recall performance for acquisition trial 1 (see **Figure 2**). By subtracting the individual's performance of AT 1 across the five trials of List A words, a better representation of the effects of repeated practice is purportedly represented.

*Learning rate* was computed by subtracting the number of words recalled in AT1 from the number of words recalled in AT5 (Lezak et al., 2004). Composite scores were computed to compare the effect of the independent variable on measures of learning. Differences in total learning, corrected total learning and learning rate scores of both groups were considered in the data analysis to assess the effects of musical verbal learning training during the learning phase.

Finally, interference in verbal learning as measured by intrusions (words not found on the target list) and repetitions (words repeated during a recall trial) were counted for the first 5 trials, at retention recall (RT) and the delayed recall trial (DRT). Contaminations (assignment of a List A or List B word to the opposing list) were counted at trials DT, RT, and DRT. Intrusions and contaminations were counted in the recognition test.

## Serial Order

Serial word order was examined through the measure of pairwise word order recall (Thaut et al., 2014). A measure of order of recall, each pair of words recalled in the order they were presented received a score of one. A total of 14 represented all 15 words recalled in order. Repeated pairs were not counted a second time, however if a single word repetition resulted in a correct word pair order that had not been counted (for e.g., "school, drum... school, parent") recall order for that word pair was counted.

## Serial Position Effects

Location of recalled words on the list was calculated by tallying the totals for each of 5 positions of list words: position A = words 1–3, position B = words 4–6, etc. Differences in serial position were compared between groups to identify potential effects of musical verbal training on primacy or recency in recall.

$$\text{Corrected Total Learning} = (\text{sum of AT 1-5}) - (5 \times \text{Sum of AT 1})$$

**FIGURE 2** | Expression for computing corrected total learning. The corrected total learning score is the sum of all words correctly recalled over the 1st 5 learning trials minus 5 times the words recalled in the 1st learning trial.

## Recognition

Recognition scoring included a tally of correctly identified List A and B words, incidence of list assignment to non-list words (intrusions) or assignment to the opposing list (contaminations).

## Procedure

The study was conducted at a private test center in the Northwestern United States. A formalized script for meeting families and child participants was used to provide greetings, and orientation and instructions about the study. The participant wore Sennheiser headphones Model HD-202 to hear a short recording to ensure satisfactory volume level, followed by an audio recording of either the control or experimental condition according to the block randomization schedule, played through an iPod Model ME178LL/A, iOS 6.1.2. The narrator voice for this test was an 11-year-old female. Words used for the memory trials and the recognition test were materials used in prior trials of the RAVLT (Schmidt, 1996).

Study participants were randomized into two groups: (1) 15-word list spoken (2) 15-word list sung [see Appendix for primary (List A) and interference (List B) words and music notation]. The control condition consisted of listening to a list of 15 unrelated words presented spoken at one per second. Immediately following hearing the list of words, the child was prompted to say back as many words as they could recall. A Sony PCM-D50 Linear PCM recorder and stereo Tascam TM-ST1 condenser microphone was used to record the participant's responses for later transcription. For the first list of words (List A), there were five acquisition trials (AT1-5). Following AT5, a distractor list of 15 words (List B) was presented (DT), and the participant was prompted to say back as many words as they could recall of the List B words. The participant was then asked to recall words from List A (RT), but without hearing them as before. After this seventh trial, a 15-min break was provided to enable a delayed recall and recognition assessment. Following the break, the participant again put on headphones and were prompted by the narrator to recall List A words (DRT). The participant was then provided with a list of 50 words including words from List A, List B and semantically and phonemically related "foil" words (Schmidt, 1996). The participant was given brief instructions to place an "A" next to the words recognized as being from List A (the list of words they heard several times) and to place a "B" next to words recognized as being from List B (the words they heard only once). Participants were instructed to ignore words they did not recognize from either List A or B.

The test condition followed the same procedures as the control condition with the exception that word lists

were presented in sung form using a novel rhythmic and melodic framework. The order and timing of words were presented just as in the control condition, but with a melodic contour and an accentuation of rhythm with two syllable words being sung as ½ second 1/8th notes (see Appendix).

### Data Analysis

An audio recording of each learning trial was coded for later review by the rater team to verify correct scoring of participant performance. Initial acquisition and learning (acquisition trials AT1-5), retention (RT, recall of primary list after presentation and recall of distractor list), and delayed recall (DRT, recall of primary list after 15 min wash-out interval) were transcribed independently by the raters. Total words recalled, repetitions, intrusions, and contaminations were tabulated for each trial, as well as serial position, and pairwise order effects. Results from hand scoring were entered into a spreadsheet for further statistical modeling and comparison of inter-rater agreement.

### Effectiveness of the Randomization

Inter-rater reliability was established through use of the Intra-Class Coefficient (ICC) statistical procedure. ICC tests were completed for dependent variables RT, DRT, T1-5PO, T7PO, T8PO, Serial position and Recognition score dependent variables. ICC comparisons of researcher and inter-rater's scores are represented in **Table 1**. Data was not found to be normally distributed so Independent Samples Mann-Whitney-U tests were used for analysis of all hypotheses and exact *p*-values were reported.

**TABLE 1** | Intra-class coefficients (ICC) for inter-rater data sets by dependent variable.

Trial	Inter-rater 1 <i>M (SD)</i>	Inter-rater 2 <i>M (SD)</i>	ICC
RT	9.9 (3.6)	9.9 (3.6)	1.000
DRT	10.1 (3.2)	10.1 (3.2)	0.998
T1-5	44.1 (11.7)	43.9 (11.7)	0.933
CTL	18.8 (7.8)	18.7 (7.4)	0.830
LR	5.9 (2.4)	5.8 (2.5)	0.982
T7PO	2.7 (3.5)	2.6 (3.6)	0.989
T1-5PO	11.1 (8.6)	10.8 (8.4)	0.991
T8PO	3.2 (3.5)	3.0 (3.2)	0.975
T1-8PO	17 (14.8)	16.9 (17.2)	0.976
Rec A	13.4 (2.1)	13.5 (2.0)	0.979
Rec B	6.8 (2.9)	6.5 (2.7)	0.793
SP1-5a	10.1 (3.2)	10.1 (3.2)	0.998
SP1-5b	9.9 (3.3)	9.9 (3.3)	1.000
SP1-5c	6.2 (3.1)	6.1 (3.1)	0.997
SP1-5d	8.3 (3.1)	8.2 (3.3)	0.989
SP1-5e	9.4 (2.9)	9.4 (3.0)	0.991

RT, Retention trial; DRT, Delayed recall trial; T1-5, Total, Trials 1-5; CTL, Corrected Total Learning; LR, Learning rate; PO, Pairwise order; Rec, Recognition scores for Lists A and B; SP, Serial position (a-e) totals for trials 1-5.

## RESULTS

Interference effects were considered by counting incidence of repetitions, intrusions and contaminations across trials. Repetitions across both groups were minimal [Trials 1-5 *M(SD)* – Spoken: 1.9 (2.8), Sung 2.8 (3.1)] and this difference was not found to be statistically significant (Mann-Whitney-U: *U* = 92.0; *p* = 0.184). Intrusions across both groups were also minimal [Trials 1-5 *M(SD)* – Spoken: 0.4 (0.6), Sung: 1.6 (2.0) and this difference was also not found to be statistically significant (Mann-Whitney-U: *U* = 83.0; *p* = 0.094)]. Contaminations were limited across both conditions [*M(SD)* – DT Spoken: 0.1 (0.3), DT Sung: 0.1 (0.3); RT Spoken: 0.2 (0.5), RT Sung: 0.0 (0.0)] and these differences were also not found to be statistically significant (Mann-Whitney-U: DT: *U* = 128.0, *p* = 1.000; RT: *U* = 112.0, *p* = 0.564). As analysis using Mann-Whitney-U tests found distributions of repetitions, intrusions and contaminations to be equally distributed in both conditions, no adjustments to raw scores were made to account for interference for any participant. As score data referenced integer (single word) counts, summary statistics were reported with one decimal point precision.

### Recall Measures

Mean recall scores for each trial, by group are provided in **Table 2**. At retention (RT), the mean of the sung word condition was 11.4 words (*SD* = 2.4) and the mean of the spoken condition was 8.4 (*SD* = 4.1) words. Analysis using the Mann-Whitney-U test found this difference at the retention stage as statistically significant (*U* = 70.5, *p* = 0.029). Analysis using Cohen's *d* found this effect size to be large (*d* = 0.893).

At the DRT, participants in the music condition still outperformed those who heard words spoken by an average of 2.6 words. Analysis using Mann-Whitney-U found this to be a statistically significant difference in performance (*U* = 73.5, *p* = 0.039). A Cohen's *d* analysis of the difference between the treatment and control group at delayed recall also found a large effect (*d* = 0.866).

**TABLE 2** | Mean recall and composite scores by group for all trials.

Trial	Spoken		Sung		Total	
	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>
AT1	16	4.8 (2.5)	16	5.3 (2.0)	32	5.1 (2.2)
AT2	16	7.6 (2.2)	16	7.9 (1.4)	32	7.8 (1.8)
AT3	16	8.8 (3.2)	16	9.9 (2.3)	32	9.3 (2.8)
AT4	16	9.9 (3.4)	16	11.4 (2.1)	32	10.6 (2.9)
AT5	16	9.8 (3.7)	16	12.2 (1.8)	32	11 (3.1)
AT1-5 total	16	41.4 (14.7)	16	46.8 (7.3)	32	44.1 (11.7)
DT (List B)	16	3.6 (1.5)	16	4.1 (1.4)	32	3.9 (1.5)
RT*	16	8.4 (4.1)	16	11.4 (2.4)	32	9.9 (3.6)
DRT*	16	8.8 (3.5)	16	11.4 (2.4)	32	10.1 (3.2)
CTL	16	17.4 (7.8)	16	20.2 (7.7)	32	18.8 (7.8)
LR*	16	4.9 (2.1)	16	6.9 (2.4)	32	5.9 (2.4)

\*Significant at *p* < 0.05; Composite scores: CTL, Corrected Total Learning; LR, Learning rate.

### Acquisition and Learning Measures

Learning was modeled through comparison of three scores: total learning, corrected total learning and learning rate. *Total learning*—the sum of all correctly recalled words across Trials AT1-5 provides a measure of overall performance over the course of multiple repetitions of the same list of words (Lezak et al., 2004). For total learning participants who listened to words sung recalled on average 5.4 more words than those in the spoken condition. This difference was not statistically significant (Mann-Whitney-U:  $U = 105$ ;  $p = 0.402$ ).

Composite scores represent quantitative approaches to parsing participant learning performance. In both composite scores of learning calculated in this study, total corrected learning and learning rate, differences were found between groups, however learning rate was the only composite score found to have a statistically significant difference between groups. On average, the participants in the song condition had a corrected learning rate of 2.8 more words than those in the spoken condition, but this positive difference was not found to be statistically significant (Mann-Whitney-U:  $U = 99.5$ ;  $p = 0.287$ ). When learning rate was considered, again a positive difference was found in favor of music, with participants recalling 2.0 words more on average in acquisition trial 5 than in acquisition trial 1 compared to participants who heard words spoken. This difference was found to be statistically significant (Mann-Whitney-U:  $U = 69.0$ ;  $p = 0.026$ ) and to have a large effect size (Cohen’s  $d = 0.887$ ). A graph of performance between conditions can be seen in **Figure 3**.

### Serial Order Effects

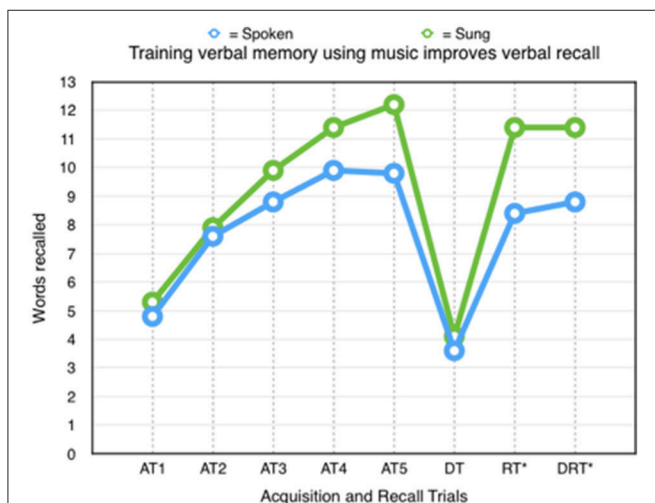
Participants in the music condition recalled more words in the order they were presented across all trials. **Table 3** provides an overview of pairwise order performance for both groups: Total word pairs recalled in order for RT and DRT, as well as totals for Trials 1-5 and total of all trials for List A (Trials 1–8).

Average pairwise word recall total for the spoken word group at Trial 7 was 1.3 words and 4.0 for the sung word group. The difference was statistically significant (Mann-Whitney-U:  $U = 71.0$ ;  $p = 0.032$ ; Cohen’s  $d = 0.827$ ). Mean pairwise word order difference between groups for Trial 8 was 3 words ( $U = 58.5$ ;  $p = 0.007$ ; Cohen’s  $d = 0.947$ ). Mean difference between groups in total of trials 1–5 was 6.6 words ( $U = 74.5$ ;  $p = 0.043$ ; Cohen’s  $d = 0.828$ ) and combining pairwise order performance across all acquisition trials of List A words (Trials 1–8), mean difference between groups was 12.2 words ( $U = 59.0$ ;  $p = 0.008$ ; Cohen’s  $d = 0.887$ ).

### Serial Position Effects

Recency and primacy effects were considered by comparing summary totals of serial position scores between groups. In serial position A (words 1–3), the average total words recalled across trials 1-5 was 10.1 for both groups. In the following positions, there were more words recalled in the music condition. Participants in the sung word group recalled a mean increased difference of 1.1 (at position B), 1.8 (at position C), 1.2 (at position D), and 1.8 (at position E) words across acquisition trials 1–5 than participants who heard words spoken. These data suggest a primacy/then recency effect in both groups with larger magnitude in the sung condition.

While these positive differences in the sung word condition were present, they did not reach statistical significance when compared to the same serial position scores of the spoken word group. As with recall data, serial position data was not normally distributed. Mann-Whitney-U tests found no significant differences between groups at any serial position [Serial Position A totals for trials 1-5(SP1-5A):  $U = 116.0$ ,  $p = 0.669$ ; SP1-5B:  $U = 111.0$ ,  $p = 0.539$ ; SP1-5C:  $U = 82.5$ ,  $p = 0.086$ ; SP1-5D:  $U = 107.5$ ,  $p = 0.445$ ; SP1-5E:  $U = 83.5$ ,  $p = 0.094$ ]. As no significant differences were found between groups in any of the positions, this statistical model does not require application of the Bonferroni correction.



**FIGURE 3 |** Comparison of groups across trials. (AT, acquisition trials; DT, distractor trials; RT, retention trial; DRT, delayed recall trial; \*Statistically significant difference).

### Recognition

Difference in raw recognition scores for both groups was minimal and influence of intrusions and contaminations was not statistically different across groups. **Table 4** provides an overview of recognition performance for both groups. Mean List A words correctly identified by participants of the spoken word group was 13 words ( $SD = 2.6$ ) while the sung word group

**TABLE 3 |** Pairwise order totals and summary scores.

Trial	Spoken		Sung		Total	
	n	M (SD)	n	M (SD)	n	M (SD)
T1-5PO*	16	7.8 (3.6)	16	14.4 (10.8)	32	11.1 (8.6)
RT-PO*	16	1.3 (1.6)	16	4.0 (4.3)	32	2.7 (3.5)
DRT-PO*	16	1.7 (1.3)	16	4.7 (4.3)	32	3.2 (3.5)
T1-8 PO*	16	10.9 (5.3)	16	23.1 (18.6)	32	17 (14.8)

\*Significant at  $p < 0.05$ ; PO, total word pairs recalled in correct serial order.

**TABLE 4** | Recognition scores for words correctly identified, intrusions, and repetitions.

Trial	Spoken		Sung		Total	
	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>n</i>	<i>M</i> ( <i>SD</i> )	<i>n</i>	<i>M</i> ( <i>SD</i> )
List A	16	13 (2.6)	16	13.8 (1.4)	32	13.4 (2.1)
A:B	16	0.9 (1.1)	16	0.4 (0.6)	32	0.7 (0.9)
A-intrusion	16	1.2 (1.4)	16	0.6 (0.6)	32	0.9 (1.1)
List B	16	7.1 (3.2)	16	6.6 (2.6)	32	6.8 (2.9)
B:A	16	0.6 (1.0)	16	0.3 (0.6)	32	0.4 (0.8)
B-intrusion	16	2.1 (4.3)	16	2.3 (2.5)	32	2.2 (3.4)

A:B, List A assigned to List B; B:A, List B assigned to List A.

participants identified 13.8 words ( $SD = 1.4$ ) (Mann-Whitney- $U$ :  $U = 108.5$ ;  $p = 0.468$ ). Mean List B words correctly identified by participants in the spoken word group was 7.1 ( $SD = 3.2$ ) while test group participants correctly identified 6.6 ( $SD = 2.6$ ) ( $U = 121$ ;  $p = 0.809$ ). While participants who listened to words sung performed marginally better in identifying List A words and spoken condition participants marginally better in identifying List B words, Mann-Whitney- $U$  tests revealed both differences to be below the threshold for statistical significance. Additionally, there was no statistically significant difference between groups for recognition intrusions (Rec-A-intrusions:  $U = 99.5$ ,  $p = 0.287$ ; Rec-B-intrusions:  $U = 95.5$ ,  $p = 0.224$ ) or contaminations (List A words assigned to List B:  $U = 89.5$ ,  $p = 0.149$ ; List B words assigned to List A:  $U = 102.0$ ,  $p = 0.341$ ).

## DISCUSSION

The purpose of this study was to compare the effects of musical mnemonics vs. spoken word in training verbal memory in children. The primary outcome of the study was shown in the retention trials where participants who listened to words sung remembered 20% more words and 17% more words at the DRT than participants who listened to words spoken. These differences were statistically significant. Results of this study provide positive initial data demonstrating benefits of musical mnemonics training of verbal memory in typically developing 9–11 year olds. Finding effective learning strategies for typically developing children is an important research goal for educational settings. A potential confound for these findings may be that treatment and control groups may not have been equivalent in IQ, as this specific measure was not considered in group assignment although all participants were from high performing school districts.

The significance of musical mnemonics training over traditional spoken verbal learning is most clearly demonstrated in the learning process after interference of the competing list (RT) and following a 15 min wash out period (DRT). These are encouraging findings for consideration of musical mnemonics in verbal learning tasks as these two points in the learning process—after interference from a distraction and after a time delay—are frequently considered measures in RAVLT reporting.

In terms of effect during the learning or acquisition phase, it appeared that the improved performance of the sung word group occurred in the later trials. Participants who listened to words sung demonstrated an only marginal mean positive difference in trials AT1 and AT2, while the differences increased with AT3 (1.19 words), AT4 (1.5 words) and especially AT5 (2.44 words). **Figure 3** shows this improved learning that was demonstrated by the participants who listened to words sung.

Markedly improved performance for test group participants in these later stages of the training phase may have contributed to the significant differences seen during the retention trial (RT) and DRT. This improved performance in the sung condition group was also present in the composite learning rate score and suggests that music's benefit to verbal memory training is more pronounced after repeated practice and the learning template or musical scaffold becomes increasingly familiar.

Similar effects were seen with serial position scores as more words from the middle to recency positions were recalled by sung group participants than by spoken group participants. Again, while these differences were not statistically significant in comparisons between groups for each position, the trend of improved performance of recently presented material is observable through comparison of means.

Previous studies of a similar test paradigm (Thaut et al., 2008, 2014) found improved pairwise word order in subjects who listened to word lists sung, in contrast to subjects who listened to the spoken condition. Results from the present study extend this observation with typically-developing children ages 9–11 as participants who listened to words sung recalled more words in the order they were presented (19% at RT and 22% at DRT) than participants who listened to words spoken. Considering the apparent superiority of musical mnemonic effects on serial order of verbal memory, clinicians should make efforts to consider this application in treatment plans for clients and patients with specific verbal memory needs.

While the results of this study are promising in the area of acquisition, recall and retention, they did not support any significant difference of repeat listening of sung verbal information on recognition. While training verbal memory through presentation of sung words was as effective as spoken word presentation in recognition, it's real value may be in recall and retention tasks where cues for retrieval are unavailable and where remembering the correct order of the information is most important.

## Implications for Clinical Practice

The effects of musical presentation on verbal memory appear to be strong, as reflected in the large effect sizes at (RT) Retention (Cohen's  $d = 0.894$ ) and (DRT) Delayed recall (Cohen's  $d = 0.893$ ). Given that this effect was identified with typically developing children, it's application for learning verbal material should be considered, especially for verbal memory tasks that require recall vs. recognition. Helping a young child remember a phone number or address may be a concern for a parent (Knott, 2017), while older children

and adults may struggle to recall passwords and academic concepts. These safety, autonomy and educational applications are but a few of the examples of how musical mnemonics can improve an individual's functioning. Additionally, these effects of improved recall and order for verbal information delivered rhythmically and melodically should be further investigated to better understand how they may be applied with special populations.

## CONCLUSION

The biggest differences between groups were seen in the area of pairwise order and may suggest a link between correct order and overall performance, as seen in previous studies. The results of this study suggest that training of verbal memory using musical mnemonics has a pronounced effect on both overall recall and serial word order recall and is superior to training verbal memory with spoken word.

## ETHICS STATEMENT

This study was carried out in accordance with the Notification for Approval of Human Subjects Research, Colorado State University Institutional Review Board, with written informed

consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The consent document was read to both the child participant and the parent. Assent and written consent was received from both child participant and parent. The protocol was approved by the Colorado State University Institutional Review Board.

## AUTHOR CONTRIBUTIONS

DK: conceived and designed the study, drafted the manuscript, tables and figures; MT: revised and provided additional content to the manuscript.

## ACKNOWLEDGMENTS

This article is based on a thesis submitted to the Academic Faculty of Colorado State University in partial fulfillment of the requirements for the degree of Master of Music and archived online (Knott, 2016). The MM thesis is the only medium this content has appeared in and its publication is in line with Colorado State University's policies. The lead author would like to thank Blythe LaGasse, Carol Seger, and Andrew Knight for advising during the design and implementation of the study, Laura Gibbons and Ann Hess for their statistical advice and Matthew Lennon for serving as inter-rater.

## REFERENCES

- Al Otaiba, S., and Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *J. Learn. Disabil.* 39, 414–431. doi: 10.1177/00222194060390050401
- Baddeley, A. D., and Hitch, G. J. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1
- Chan, A. S., Ho, Y. C., and Cheung, M. C. (1998). Music training improves verbal memory. *Nature* 396, 128–128.
- Cirigliano, M. M. (2013). Musical mnemonics in health science: a first look. *Med. Teach.* 35, e1020–e1026. doi: 10.3109/0142159X.2012.733042
- Claussen, D. W., and Thaut, M. H. (1997). Music as a mnemonic device for children with learning disabilities. *Can. J. Music Ther.* 5, 55–66.
- Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Dir. Psychol. Sci.* 18, 275–279. doi: 10.1111/j.1467-8721.2009.01651.x
- Dehn, M. J. (2010). *Long-Term Memory Problems in Children And Adolescents: Assessment, Intervention, And Effective Instruction*. Hoboken, NJ: John Wiley and Sons.
- Deutsch, D. (1982). "Organizational processes in music," in *Music, Mind, and Brain*, ed M. Clynes (Boston, MA: Springer), 119–136.
- Forrester, G., and Geffen, G. (1991). Performance measures of 7- to 15-year-old children on the auditory verbal learning test. *Clin. Neuropsychol.* 5, 345–359. doi: 10.1080/13854049108404102
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608. doi: 10.1093/brain/awl247
- George, E. M., and Coch, D. (2011). Music training and working memory: an ERP study. *Neuropsychologia* 49, 1083–1094. doi: 10.1016/j.neuropsychologia.2011.02.001
- Gfeller, K. E. (1983). Musical mnemonics as an aid to retention with normal and learning disabled students. *J. Music Ther.* 20, 179–189. doi: 10.1093/jmt/20.4.179
- Habib, M., and Besson, M. (2009). What do music training and musical experience teach us about brain plasticity? *Music Percept.* 26, 279–285. doi: 10.1525/mp.2009.26.3.279
- Ho, Y. C., Cheung, M. C., and Chan, A. S. (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology* 17:439. doi: 10.1037/0894-4105.17.3.439
- Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Peterson, R. C., Kokmen, E., et al. (1992). Mayo's older Americans normative studies: updated AVLT norms for ages 56-97. *Clin. Neuropsychol.* 6, 83–104.
- Knott, D. (2016). *Immediate Effects of Training with Musical Mnemonics on Verbal Memory in Children*. Master's thesis, Available from ProQuest Dissertations and Theses database. (UMI No. 10241676).
- Knott, D. (2017). Musical mnemonics training: proposed mechanisms and case example with acquired brain injury. *Music Ther. Perspect.* 35, 23–29. doi: 10.1093/mtp/miv016
- Lezak, M. D. (1983). *Neuropsychological Assessment, 2nd Edn*. New York, NY: Oxford University Press.
- Lezak, M. D., Howieson, D. B., and Loring, D. W. (2004). *Neuropsychological Assessment, 4th Edn*. New York, NY: Oxford University Press.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63:81. doi: 10.1037/h0043158
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Peterson, D. A., and Thaut, M. H. (2007). Music increases frontal EEG coherence during verbal learning. *Neurosci. Lett.* 412, 217–221. doi: 10.1016/j.neulet.2006.10.057



- Schmidt, M. (1996). *Key Auditory Verbal Learning Test: A Handbook*. Los Angeles, CA: Western Psychological Services.
- Thaut, M. H., Peterson, D. A., McIntosh, G. C., and Hoemberg, V. (2014). Music mnemonics aid verbal memory and induce learning-related brain plasticity in multiple sclerosis. *Front. Hum. Neurosci.* 8:395. doi: 10.3389/fnhum.2014.00395
- Thaut, M. H., Peterson, D. A., Sena, K. M., and McIntosh, G. C. (2008). Musical structure facilitates verbal learning in multiple sclerosis. *Music Percept.* 25, 325–330. doi: 10.1525/mp.2008.25.4.325
- Wallace, W. T. (1994). Memory for music: effect of melody on recall of text. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1471.
- Wan, C. Y., and Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *Neuroscientist* 16, 566–577. doi: 10.1177/1073858410377805
- Wolfe, D. E., and Hom, C. (1993). Use of melodies as structural prompts for learning and retention of sequential verbal information by preschool students. *J. Music Ther.* 30, 100–118. doi: 10.1093/jmt/30.2.100

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Knott and Thaut. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## APPENDIX

### Primary and Interference Lists Music Notation

**Primary (List A) melody**

♩ = 60

drum cur-tain bell cof-fee school par-ent moon gar-den hat fir-mer nose tur-key col-or house ri-ver

**Interference (List B) Melody**

♩ = 60

desk ran-ger bird shoe stove moun-tain glass-es tow-el cloud boat lamb gum pen-cil church fish

**FIGURE A1** | Primary and interference lists music notation. Words, melodic and rhythmic notation used in Primary and Interference lists.