https://doi.org/10.24839/2325-7342.JN22.2.141

Lyrical Memory: Mnemonic Effects of Music for Musicians and Nonmusicians

Sarah N. Lummis, Jennifer A. McCabe*, Abigail L. Sickles, Rebecca A. Byler,
Sarah A. Hochberg, Sarah E. Eckart, and Corinne E. Kahler
Goucher College

ABSTRACT. The goal of this study was to examine the mnemonic effects of music on lyric recall, and to determine the role of musical expertise in the degree of memory benefit from information presented in a song. The experiment used a mixed-factor design to compare recall performance in music and nonmusic presentations, in addition to comparing performance of musicians and nonmusicians. Participants were presented with a set of lyrics in one of two encoding conditions: music (delivered as part of a song) or nonmusic (read aloud without music), followed by a free recall test. This process repeated 3 times, with a distractor task before the third recall test. One week later, participants were given an identical series of encoding and recall tasks in the other encoding condition, followed by a recall test for lyrics presented in the first session. Results showed significantly higher recall in the music condition during Session 1 as measured by words (verbatim), lines (gist), and clusters (chunking), $ps < .05$; for delayed recall, there was a music advantage as measured by words and clusters, $ps < .05$. Musicians showed significantly higher 1-week-delayed word and line recall, regardless of encoding condition, $ps < .05$. Several significant differences were found in relation to task load, suggesting that music-based learning may affect subjective experience, specifically task success and time pressure; further, musicians reported lower mental activity required when learning through music. Further applications of the study are discussed in an educational context.

Learning through song is a commonly used strategy—after all, many children are taught lyrics with melody to learn the alphabet, days of the week, the 50 states, and more. Music-based learning can also be effective for more dense material and difficult facts. For example, there are songs to support the learning of statistics, the periodic table of elements, and foreign languages. The current study replicated and extended prior research on learning through music, with the goal of establishing more definitively the benefit of music as a mnemonic tool for college-aged younger adults.

Mnemonics are memory strategies that organize to-be-learned information in ways that enable encoding and facilitate later retrieval. Various types of verbal mnemonics (e.g., acronyms, acrostics) and visual-spatial mnemonics (e.g., keyword, method of loci, pegword) have been shown to improve memory for longer term retention of information (Bellezza, 1996). Music is promising as an effective mnemonic because it offers unique chunking and associative strategies based on rhythm, melody, and pacing. Relevant to pacing, temporal or time-based aspects of songs have been considered especially helpful when organizing information in memory (see Yalch, 1991, for a review). Additionally, evidence that people integrate the tune and the lyrics in the resulting mental representation is consistent with the idea that musical elements

*Faculty mentor

support memory for lyrics (Morrongiello & Roes, 1990; Serafine, Crowder, & Repp, 1984).

Prior research on the use of music as a mnemonic has generally supported the current hypothesis that setting lyrics to music will improve recall. Several studies have focused on music and memory in children, finding an advantage for music as a mnemonic aid (Campabello, De Carlo, O'Neil, & Vacek, 2002; Gingold & Abravanel, 1987). With regard to the population of interest for the current study, college-aged adults, Chazin and Neuschatz (1990) compared the effectiveness of learning scientific information through a familiar melody versus through lecture. Results showed that song-based information produced a higher recall over lecture-based information (see also commentary by Scruggs & Brigham, 1991). In another study that utilized familiar melodies, Rainey and Larsen (2002) found no effect of music on initial learning, but did find faster relearning in participants who heard the sung version of the information.

However, other research has failed to show a memory impact for musical presentation of information. Moore, Peterson, O'Shea, McIntosh, and Thaut (2008) studied music as a mnemonic in people with multiple sclerosis and healthy controls. The results did not show significant differences in recognition memory between groups or between music and spoken conditions. Moore et al. discussed that these null results could have been due to differences in the format of learning and testing conditions, and also insufficient time for encoding. Deason et al. (2012) examined recognition memory performance in patients with Alzheimer's and healthy older adults. Both groups listened to lyrics accompanied by music and lyrics spoken without music, while having the lyrics in front of them. There was no significant difference in memory performance, including at a 1-week delay, between sung versus spoken lyrics in the healthy older adult group. Although on the surface these studies do not appear support our hypothesis about music benefitting memory, it is important to note the differences between their studies and the current experiment—namely, the focus on comparing preexisting participant groups, and the choice of stimuli and procedures.

Some research has even suggested superior memory for spoken words compared to song conditions. For example, Calvert and Billingsley (1998) compared children's memory of their phone numbers presented in a song versus spoken condition. The results indicated that the children recited the spoken condition better than the song. Researchers suggested that songs offer children stimuli that were too challenging to their minimal attention and processing resources, a stark contrast to our study that examined adults with better ability to handle task load.

In addition to examining the effectiveness of music as a mnemonic device, the current study examined whether music "experts" (i.e., musicians) can use the musical structure of a song to better or more efficiently memorize lyrics compared to nonmusicians, perhaps due to enhanced chunking. Some research has suggested an advantage for musicians when learning from musical stimuli, whereas other studies have failed to show an advantage for those with musical training. Ginsborg (2007) had singers perform an a capella song from memory after they had intentionally learned and memorized the lyrics and melody, both independently and together. Musicians performed more accurately and articulately than participants who had less musical expertise, but this happened only when they memorized the lyrics and melody together. Ginsborg argued that lyrics and melody are recollected together, such that recalling one allows the recall of the other. These findings support the hypothesis that memorizing lyrics and melody together can be a successful strategy, especially for musicians.

Furthermore, Silverman (2007) studied the effect of paired pitch, rhythm, and speech on college students' memory, using sequential digit recall as the dependent measure. Participants listened to four counterbalanced conditions, each consisting of nine randomized one-syllable digits paired with speech, pitch, rhythm, and the combination of pitch and rhythm. Participants recalled digits from the rhythm condition most precisely, and recalled digits from the speech and pitch only conditions the least precisely. Most important for the current study, the music majors scored significantly higher than the nonmusic majors.

Conversely, some research has shown no effect of music on memory recall for musicians. In Racette and Peretz's (2007) study, participants learned an unfamiliar song in three conditions: *sung-sung* (the song being learned was sung, and the response was sung), *sung-spoken* (the response was spoken), and *divided-spoken* (the presented lyrics were accompanied by music and the response were spoken). Fewer words were recalled when singing than when speaking in both the musician and nonmusician groups. Taken together, these conflicting results speak to the need for further research on the effect of music on memory recall, specifically in musicians.

As a means of understanding why musical experts may experience a stronger benefit of music as a mnemonic, and to better understand the broad theoretical frame regarding the role of musicians versus nonmusicians in the current study, it is helpful to consider classic and contemporary research on chess players, specifically focusing on the role of chunking as a hallmark of expertise. Chase and Simon (1973) found that master chess players more readily memorized a midgame configuration of chess pieces compared to chess novices; due to their extensive prior knowledge about chess, they perceived and encoded groups (or chunks) of pieces in meaningful sets. However, this memory advantage did not appear when memorizing chess pieces in random placements. Thus, experts may have a specific advantage in meaningfully chunking information in their domain of expertise. More recently, Bilalić, Langer, Erb, and Grodd (2010) showed that it took expert chess players half the time to count knights and bishops in a chess task compared with novices, particularly in the condition that was identical to a midgame scenario; yet contrary to Chase and Simon's result, experts still had an advantage over novices for randomly placed positions in the chess task. Other researchers have used neurophysiological measures to look at chunking. Amidzic, Riehle, and Elbert (2006) found that the location of brain activity is different in experts, who show activation that is consistent with the retrieval of chunks, compared to nonexperts, who show activation in areas for memory formation. The results from these expertise and chunking studies support the general idea that musicians (as experts in the domain of music) may show a larger advantage for music-based encoding compared to nonmusicians.

The current study offered a partial replication and extension of two prior studies that used existing but unfamiliar songs as stimuli, and that found superior memory for musical encoding conditions in younger adults. First, Wallace (1994) presented verses of 80 to 85 words, with and without accompanying melody; results showed that, as long as the stimulus was repeated sufficiently for initial learning, there was an advantage for sung lyrics over read-aloud lyrics. Second, McElhinney and Annett (1996) compared the impact of hearing sung lyrics to read-aloud lyrics; each presentation was followed by a lyric recall test. Results showed that sung lyrics were recalled at a higher level in the second and third rounds of recall, and also that recall patterns for songs showed more evidence of chunking (i.e., recall in groups or phrases versus single words).

The purpose of the present study was to investigate music as a mnemonic device using elements of experimental design to improve internal validity (e.g., use of two counterbalanced songs/lyrics as stimuli, delayed recall tests). The two main research questions were as follows: (a) Is memory recall better for lyrics set to music compared to lyrics read without music? (b) Do musicians and nonmusicians benefit differently from music as a mnemonic device? Although prior research on this topic has been mixed, we hypothesized that the recall of musical lyrics would overall be higher than the recall of nonmusical lyrics at each time point. Further, we predicted an interaction between encoding condition and musicianship, such that musicians would have a specific advantage in recalling lyrics set to music.

In addition to analyzing the objective measure of recall, we also aimed to focus on subjective measures of task experience. We included questions about level of task load experienced during the recall task, with the hypothesis that learning lyrics with music may impact phenomenological perceptions of mental activity, time pressure, and/or task success; and that musicians may experience music-based learning differently than nonmusicians.

Method

Design

The experiment was a mixed-factor 2 (encoding condition: music, nonmusic) x 2 (expertise: musician, nonmusician) x 3 (Time: Recall 1, Recall 2, Recall 3) design. Encoding condition and time were manipulated within-subjects, and expertise was a between-subjects factor.

Participants

One hundred fifteen undergraduate students were recruited from various psychology courses and music groups at a private liberal arts college. Twenty students failed to return for the second session, and one student indicated familiarity with recall stimuli used; this resulted in the omission of 21 participants from the data set, leaving a final sample size of $n = 94$. Students were offered extra credit in some psychology courses, or entry in a gift card raffle in exchange for participation. The sample had a mean age of 19.79 ($SD = 1.39$) and included 24.2% men, 72.6% women, and 2.1% other/preferred not to disclose.

Musicianship of participants was determined based on a survey about past musical experience, resulting in a reasonably equal comparison of *n* = 44 (47%) musicians and *n* = 50 (53%) nonmusicians. Using responses from this survey, researchers classified each participant as a musician or as a nonmusician. Two researchers separately reviewed each survey and decided whether each participant would fit the description of a musician in the context of the current study, based on the extent of their musical training. A third researcher mediated any discrepancies between the judgments of the two researchers.

Materials

Two sets of lyrics were chosen for the study from pre-existing songs. The musical stimuli, presented to participants in the music condition, consisted of portions from two songs that were delivered in a moderate musical pacing (mean BPM = 110) and were both sung by men. Portions of the songs were chosen to exclude any repeated choruses or lines, so there was very little lyrical repetition. In addition, both portions portray a narrative storyline, in order to open the potential of examining not only word recall, but a more general idea recall as well. The songs were "Taxi" by Harry Chapin and "Cold Missouri Waters" by Cry Cry Cry. The selection of lyrics from "Taxi (Lyric A) included 147 words and 21 lines; the music clip was 1 minute 54 seconds long, and the nonmusic clip was 1 minute 3 seconds long. The selection of lyrics from "Cold Missouri Waters" (Lyric B) included 148 words and 17 lines; the music clip was 1 minute 53 seconds long, and the nonmusic clip was 1 minute 8 seconds long. In the nonmusic condition, participants were presented with a prerecorded reading of the lyrics. Each reading was done in a male voice, and each followed the pace of the original song. Participants were asked about their prior familiarity with the songs at the conclusion of each session; as noted above, the one participant who indicated familiarity was not included in the data set.

In addition to a sheet containing the printed lyrics, each participant was given a testing packet, which contained a blank sheet of paper for each free recall test, and an inventory to measure task load. The inventory contained six modified questions from the NASA-Task Load Index (Hart & Staveland, 1988), which were answered on a 7-point Likert-type scale. Questions were chosen and modified from the original scale to pertain to the task of lyric encoding and recall (see Table 1 for the modified questions). This inventory allows for items to be analyzed individually, each as a separate indicator of task load.

Testing packets for Session 2 contained a survey with demographic and musical background questions. In addition to asking participants to indicate whether they self-identify as musicians, the musical experience survey inquired about the participants' level, type, and length (number of years) of musical experience. Finally, to collect further data regarding musicianship, the survey asked participants two questions on a 5-point Likert-type scale ranging from 1 (*definitely no*) to 5 (*definitely yes*): "Can you sight-read an easy piece of music?" and "If a melody was sung to you, could you sing it back accurately?"

Procedure

Following institutional review board approval (2014124R), testing sessions were conducted in groups of 1 to 10 participants. Counterbalancing was achieved by using two different lyric sets and two different encoding conditions (music and nonmusic), forming four possible conditions: musical presentation of Lyric A (*MA*), nonmusical (spoken) presentation of Lyric A (*NMA*), musical presentation of Lyric B (*MB*), nonmusical (spoken) of Lyric B (*NMB*). Based on the assigned condition of Session 1, participants received the alternate pairings of each of the variables in Session 2. Thus, the four possible counterbalancing conditions are as follows (given in the format of Session 1-Session 2): *MA-NMB* (*n* = 22), *MB-NMA* (*n* = 23), *NMA-MB* (*n* = 23), and *NMB-MA* (*n* = 26). Groups of participants were initially assigned to a counterbalancing condition with the goal of achieving an equal distribution of participants in each condition. Results showed that there were no differences in memory recall performance across counterbalancing orders, so we will not discuss this variable further.

Upon arrival to the testing room, participants read and signed the informed consent form provided, then were each given a testing packet and a sheet of paper containing the lyrics of the audio clip to be presented during that session. They were told that they would listen to an audio clip three separate times, and that they would be asked to recall as much as they could after each presentation. Participants were asked to silently read along with the lyrics as they heard them. Upon the conclusion of the audio clip, they put the lyric sheet away, turned to a blank page in their testing packets, and wrote down as many of the lyrics as they could remember including words, phrases, or general ideas, in a 3-minute period. The audio clip was then repeated, and they again completed the free recall test for 3 minutes. After the audio clip was played a third and final time, participants were given a 5-minute distractor task (a paper packet with several different puzzles) before completing the third free recall test. Participants then turned to the final page in their testing packet to complete the task load questionnaire (see Table 1 for items).

The second session, exactly 1 week later, was identical to the first, except for the encoding condition and lyric set (see counterbalancing description above), and the fourth free recall test at the end of the session. The fourth recall test was given to test longer term memory for the lyrics presented in the first session. Thus, approximately half of the participants performed a 1-week-delayed recall task from material in the music condition, and the other half recalled from the nonmusic condition. Unlike the other recall variables, which were manipulated within-subjects, the measure of delayed recall of lyrics from Session 1 was a between-subjects variable. Finally, at the conclusion of Session 2, participants were asked to complete a questionnaire to indicate demographic information and the nature of their musical experience (see Figure 1 for an overview of the study design). In summary, participants completed three recalls of one set of lyrics during one session; then, one week later, they completed three recalls of another set of lyrics, along with one recall for the lyrics from the previous week.

Free recall tests were scored for exact (verbatim) word recall and more leniently for line recall. Line recall was implemented to measure memory for general ideas, and used a point system in which one point was awarded for every line of the lyrics for which the scorer judged that the participant accurately expresses the idea or gist, regardless of exact word recall. Final proportion scores for both words and lines were calculated by dividing the units (words or lines) recalled by the maximum number of units in the given set of lyrics. Each response was scored separately by two different researchers, and any discrepancies were mediated by a third researcher, in order to ensure consistency.

Scoring for evidence of chunking was also done by a point system, based on *clusters,* which are operationally defined as a group of three to five words in a participant's response that were recalled

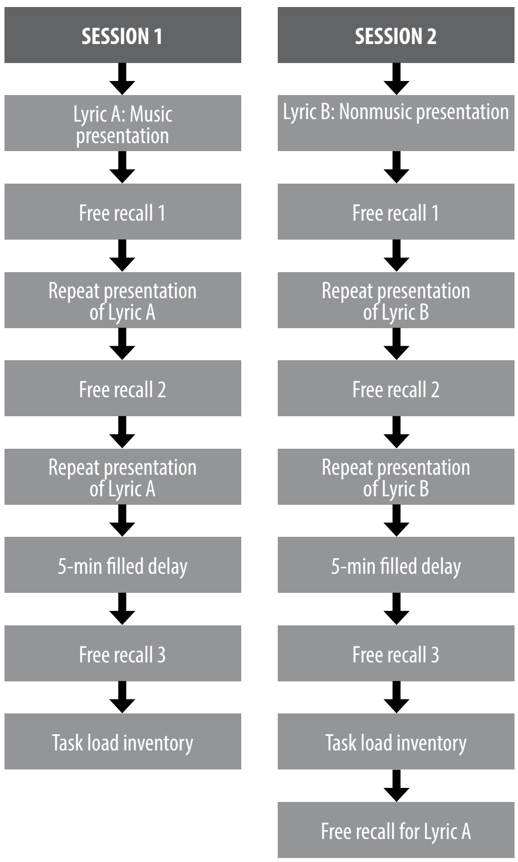
TABLE 1

Means and Standard Deviations for Task Load Items in Music and Nonmusic Encoding Conditions for Musicians and Nonmusicians

	Music Encoding			Nonmusic Encoding		
Question	Musician	Nonmusician	Overall	Musician	Nonmusician	Overall
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
1 – Mental	4.83 (1.39)	5.08 (1.24)	4.97 (1.31)	5.06 (1.24)	4.80 (1.06)	4.93 (1.15)
2 – Time	4.80 (1.71)	4.51 (1.63)	4.68 (1.67)	4.50 (1.89)	4.31 (1.89)	4.38 (1.81)
3 – Success	4.49 (1.12)	3.94 (1.24)	4.19 (1.21)	3.96 (1.01)	3.67 (1.20)	3.80 (1.11)
4 – Hard	4.54 (1.47)	4.66 (1.33)	4.60 (1.34)	4.71 (1.53)	4.37 (1.42)	4.54 (1.48)
5 – Tense	3.47 (1.40)	4.00 (1.59)	3.75 (1.52)	3.77 (1.49)	3.75 (1.60)	3.79 (1.56)
6 – Frustrated	3.28 (1.62)	3.34 (1.56)	3.32 (1.59)	3.36 (1.56)	3.24 (1.47)	3.28 (1.50)

Note. Task load questions were measured on a 7-point scale from low to high, using a modified version of the NASA-TLX (Hart & Staveland, 1988). Questions included (1; *Mental*) "How much mental activity was required for the tasks?" (2; *Time*) "How much time pressure did you feel due to pacing of the tasks?" (3; *Success*) "How successful do you think you were in accomplishing the goals of the tasks set by the experimenter?" (4; *Hard*) "How hard did *you* have to work to accomplish your level of performance?" (5; *Tense*) "How tense versus relaxed did you feel during the tasks?" and (6; *Frustrated*) "What was your level of frustration during the tasks?"

FIGURE 1



Overview of study design. Counterbalancing was achieved by alternating which lyric set and which type of presentation was used first. This figure represents one of four possible counterbalancing conditions (i.e., music condition for Lyric A first, nonmusic condition for Lyric B second); otherwise the testing procedure presented above was identical for all participants. Session 2 occurred 1 week after Session 1.

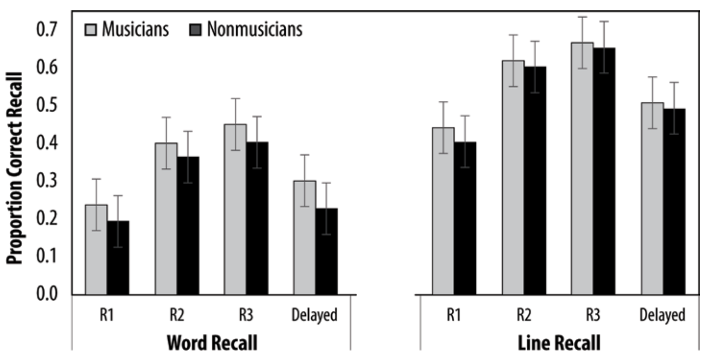
TABLE 2

Means and Standard Deviations for Recall in Music and Nonmusic Encoding Conditions for Musicians and Nonmusicians

	Music Encoding			Nonmusic Encoding		
	Musician	Nonmusician	Overall	Musician	Nonmusician	Overall
Recall	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Word R1	0.25 (0.12)	0.23 (0.12)	0.24 (0.12)	0.21 (0.11)	0.18 (0.10)	0.19 (0.10)
Word R2	0.42 (0.15)	0.38 (0.15)	0.40 (0.15)	0.38 (0.14)	0.35 (0.12)	0.36 (0.13)
Word R3	0.48 (0.15)	0.42 (0.16)	0.45 (0.16)	0.43 (0.13)	0.38 (0.16)	0.40 (0.15)
Overall	0.38 (0.14)	0.34 (0.14)	0.36 (0.17)	0.30 (0.13)	0.34 (0.12)	0.32 (0.16)
Delayed Word	0.38 (0.18)	0.24 (0.14)	0.31 (0.17)	0.24 (0.16)	0.22 (0.16)	0.23 (0.16)
Line R1	0.44 (0.16)	0.44 (0.17)	0.44 (0.16)	0.42 (0.15)	0.38 (0.14)	0.40 (0.14)
Line R2	0.62 (0.13)	0.61 (0.16)	0.62 (0.14)	0.61 (0.15)	0.59 (0.15)	0.60 (0.15)
Line R3	0.68 (0.16)	0.65 (0.16)	0.67 (0.16)	0.68 (0.13)	0.63 (0.18)	0.65 (0.16)
Overall	0.58 (0.15)	0.57 (0.16)	0.57 (0.18)	0.57 (0.14)	0.53 (0.14)	0.55 (0.19)
Delayed Line	0.59 (0.20)	0.52 (0.18)	0.55 (0.19)	0.44 (0.20)	0.47 (0.23)	0.46 (0.22)

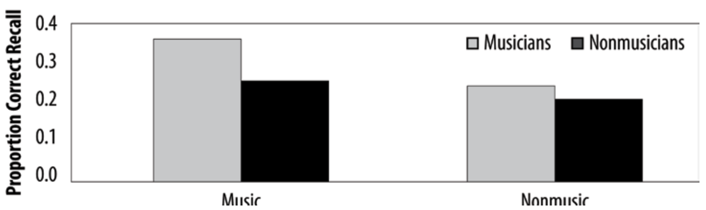
Note. Recall 1 (R1) and Recall 2 (R2) each occurred after an opportunity to encode the lyrics while listening and looking at them. Recall 3 (R3) occurred after a filled delay. Delayed recall occurred one week after the initial session. Word recall represents verbatim memory for lyrics, and line recall represents memory for general ideas in lines of the song.

FIGURE 2



Word and line recall in music and nonmusic encoding conditions. "Delayed" represents 1-week-delay recall administered during the second testing session. Error bars represent standard errors.

FIGURE 3



Word recall scores for musicians and nonmusicians for both encoding conditions in the 1-week-delayed recall.

in the same order in which they were presented in the original stimulus (Meijs, Hurks, Rozendaal, & Jolles, 2013). For example, with the original lyrics reading, "she just looked out the window…" a participant would receive one point for "looked out the window." Given the similar number of words in each set of lyrics, we did not compute proportion correct scores for Lyric A and Lyric B; instead, we report absolute number of clusters below.

Results

Musicianship Ratings

Pearson correlation analyses among pairs of items on the musicianship survey (see details above in *Method* section) showed strong positive correlations (*ps* < .001) for all variable pairs. Given that researcher-determined and self-reported musicianship ratings were strongly correlated, $r(92) = .730$, *ps* < .001, the researcher-determined ratings are used to classify musicianship as a dichotomous variable (i.e., musician, nonmusician) in the remainder of the analyses. Interrater reliability was determined using Cohen's κ, which showed an almost perfect agreement between researchers' judgments of musicianship, κ = .914, p < .001.

Lyric Recall

The alpha level was set at p = .05. We report η_p^2 as a measure of effect size: $\eta_p^2 = .01$ represents a small effect; $\eta_p^2 = .06$, a medium effect; and $\eta_p^2 = .14$, a large effect (Cohen, 1988). An outlier analysis of each recall variable was conducted based on converting raw accuracy scores to *z* scores, then evaluating against the commonly used criterion of $z = 3.29$. Scores that exceed this criterion would be considered outliers, falling in the top or bottom 0.1% of the distribution. These analyses revealed no outlying scores in any of the variables.

Proportion of words recalled. A 2 x 3 x 2 mixed-factor Analysis of Variance (ANOVA) was conducted to analyze the effect of encoding condition, time, and musicianship on proportion of words recalled (see Table 2 and Figure 2). Results showed a significant increase in word recall across each of three time periods, $F(2, 188) = 438.65$, $p < .001$, $\eta_p^2 = .828$, with all contrasts being significantly different, *ps*<.001 (see Table 2 for all descriptive statistics). Word recall was also significantly different between encoding conditions, $F(1, 92) = 15.86$, $p < .001$, $\eta_p^2 = .148$, with higher scores in the music condition ($M = 36$, $SD = 0.16$) than in the nonmusic condition ($M = 0.32$, $SD = 0.15$). In regard to musicianship, word recall for musicians

(M = 0.36, SD = 0.17) was not significantly higher than that for nonmusicians (M = 0.32, SD = 0.16), $F(1,92) = 2.79$, $p = .099$, $\eta_p^2 = .030$. No interactions were found between any of the variables.

Next, to examine the 1-week-delayed recall test for lyrics presented in the first session, a 2 (encoding condition) x 2 (musicianship) ANOVA was conducted on word recall. A significant difference was found in relation to encoding condition, $F(1, 92) = 4.96$, $p = .029$, $\eta_p^2 = .058$, with higher scores in the music condition (M = 0.30, SD = 0.18) than in the nonmusic condition (M = 0.23, SD = 0.16). A significant difference was also found in relation to musicianship, $F(1, 92) = 5.82$, $p = .018$, $\eta_p^2 = .067$, with musicians having higher scores (M = .31, SD = .17) than nonmusicians (M = 0.23, SD = 0.16). There was no interaction, $F(1, 92) = 2.65$, $p = .107$, $\eta_p^2 = .032$ (see Figure 3).

Proportion of lines recalled. Similar results were found in regard to line recall (see Table 2 and Figure 2). A 2 x 3 x 2 ANOVA showed a significant increase in line recall scores in relation to time of recall, $F(2, 188) = 227.23$, $p < .001$, $\eta_p^2 = .710$, with all contrasts being significantly different, ps <.001. A significant difference in line recall was also found in relation to encoding condition, $F(1, 92) = 4.09$, $p = .046$, $\eta_p^2 = .043$, with higher scores in the music condition (M = 0.58, SD = 0.14) than in the nonmusic condition (M = 0.55, SD = 0.18). Line recall for musicians (M = 0.43, SD = 0.16) was not significantly higher than that for nonmusicians (M = 0.41, SD = 0.16), $F(1, 92) = 1.25$, $p = .267$, $\eta_p^2 = .014$, and no interactions were found between any of the variables.

Turning to results for 1-week-delayed line recall, a significant difference was found in relation to musicianship, $F(1, 92) = 4.25$, $p = .042$, $\eta_p^2 = .050$, with musicians showing significantly higher scores (M = 0.55, SD = 0.19) than nonmusicians (M = 0.46, SD = 0.22). No significant differences were found in relation to encoding condition, $F(1, 92) = 0.23$, $p = .631$, $\eta_p^2 = .003$; further, there was no interaction.

Cluster scores. Cluster scores produced similar patterns to that of line and word recall. Scores significantly increased with each recall, $F(2, 188) = 217.92$, $p < .001$, $\eta_p^2 = .710$, with all contrasts being significantly different, ps <.001. In addition, there was a significant difference in relation to encoding condition, $F(1, 92) = 16.32$, $p < .001$, $\eta_p^2 = .155$, with higher scores in the music condition (M = 11.80, SD = 7.76) than in the nonmusic condition (M = 9.71, SD = 6.35). For the 1-week-delayed recall, scores were also significantly higher in the music (M = 9.03, SD = 7.34) than in the nonmusic condition (M = 5.42, SD = 4.80), $F(1, 92) = 7.40$, $p = .008$, $\eta_p^2 = .086$. No significant differences were found in relation to musicianship. Not surprisingly, cluster scores were positively correlated with word and line recall scores in the corresponding encoding conditions, $ps < .001$.

Task Load

A series of 2 (encoding condition) x 2 (musicianship) mixed-factor ANOVAs was used for the task load questions (see Table 1). For the question, "How much mental activity was required for the tasks," there were no main effects ($ps > .05$), but there was a significant interaction, $F(1, 92) = 4.22$, $p = .043$, $\eta_p^2 = .045$. This interaction was driven by musicians reporting higher mental activity required for the task in the nonmusic condition, whereas nonmusicians reported higher mental activity in the music condition.

Analyses of the question, "How much time pressure did you feel due to pacing of the tasks," showed a significant main effect of encoding condition, $F(1, 92) = 4.09$, $p = .046$, $\eta_p^2 = .043$; participants reported higher time pressure in the music condition (M = 4.69, SD = 1.66) than in the nonmusic condition (M = 4.38, SD = 1.47). No other effects were significant ($ps > .05$).

For the question, "How successful do you think you were in accomplishing the goals of the tasks set by the experimenter," there was a main effect of encoding condition, $F(1, 92) = 10.34$, $p = .002$, $\eta_p^2 = .102$. Participants reported significantly higher perceived success in the music condition (M = 4.19, SD = 1.21) compared to the nonmusic condition (M = 3.81, SD = 1.11). We found no significant findings in relation to the remainder of the task load questions ($ps > .05$).

Discussion

The current study examined the effect of music on recall. Participants listened to lyrics set to music (or read aloud without music) three times, each followed by a free recall test, with a 5-minute delay period preceding the third test. One week later, participants repeated the procedure in the other encoding condition (music, nonmusic). They then completed a free recall test on the lyrics learned in the testing session 1 week prior. Recall was measured by percentage of words recalled, percentage of lines recalled, and number of clusters in each response. After both testing sessions, participants completed a task load inventory to measure the subjective mental demands of the recall tasks.

As predicted, music-based encoding produced overall higher recall than nonmusic encoding. Participants showed higher word recall, line recall, and cluster scores across all three same-session recall tests when learning lyrics as part of the original song; this pattern was also evident on the 1-week-delayed recall test, for words and clusters. Music-based encoding, therefore, showed an advantage in long-term memory. These results not only provide support for the original hypothesis, but they are consistent with the previous literature regarding the effectiveness of music as a mnemonic device to improve recall of information (Campabello et al., 2002; Chazin & Neuschatz, 1990; Gingold & Abravanel, 1987). The finding that music encoding had an advantage in recall of specific verbatim information, ideas or gist, *and* chunks of information, has implications for educational settings: Learning information set to music may be useful in situations ranging from detailed memorization to the learning of general concepts and ideas.

The hypothesis stated that musical expertise may play a role in learning through song, based on findings of a music-based memory advantage for musicians (Ginsborg, 2007; Silverman, 2007) and extrapolated from the idea that experts have a chunking advantage in their domain of expertise (Chase and Simon, 1973). Contrary to this hypothesis, in the current study, there was no unique recall advantage for musicians in the music condition. Interestingly, however, musicians had significantly higher word and line recall scores on the 1-week-delayed recall test, regardless of encoding condition. Beyond the scope of this study, it is possible that the superior scores of the musicians reflected an overall advantage in general mental abilities as a result of musical training, and/or due to pre-existing characteristics (e.g., working memory capacity, verbal recall; Brandler & Rammsayer, 2003).

Results from the task load inventory provided a lens into the subjective experience of learning with and without music. The modified NASA-Task Load Index (Hart & Staveland, 1988) was administered at the end of both testing sessions to examine whether music-based encoding would change the perception of mental activity, effort, time, or task success. An interesting pattern was found in regard to the question, "How much mental activity was required for the tasks?" Whereas musicians reported higher mental activity required for the task in the nonmusic condition, nonmusicians reported higher mental activity required for the task in the music condition. Because this trend does not parallel the objective recall scores across conditions and participant groups, it is perhaps a reflection of an increased comfort level for musicians when experiencing a memory task in the context of their domain of expertise.

Additionally, participants (regardless of musical expertise) reported feeling higher time pressure from the task in the music condition, compared to the nonmusic condition. It can be speculated that the explicitly rhythmic and paced nature of songs, with accompanying melody and singing, might have caused participants to feel more rushed during the encoding task, in comparison to hearing the lyrics spoken aloud without music. It is also possible that these ratings were affected by the level of recall on the subsequent memory tests; given that the task load questions were administered at the very end of the session, the fact that participants were able to recall more words in the music condition might have contributed to their retroactive perceptions of a higher time constraint on encoding.

Finally, participants reported significantly higher perceived success in the music condition compared to the nonmusic condition. This is in accordance with the finding that participants had higher recall scores in the music condition. As such, participants seemed to be metacognitively aware that they were learning more of the lyrics when presented as part of a song.

The results overall suggest that music can aid in easier encoding and subsequent recall of information and that musicians may have an overall advantage in delayed recall of both sung and read lyrics. Results also suggest that music-based learning may cause participants to experience higher time pressure and also higher task success, and that musicians may experience a lower mental activity load when learning with music, as compared to nonmusicians.

Limitations and Future Directions

Several limitations should be taken into account when considering the results of the current study. To begin, even with equivalent pacing of songs and spoken lyrics, musical structure does add multiple layers of information to the encoding task, and the current design did not allow an examination of how those elements impact memory. It would be interesting to include as an additional comparison condition a set of lyrics that are rhythmically paced but not musical with regard to pitch.

In addition, seven participants' data had to be omitted from the 1-week-delay recall, due to the researchers' knowledge that the participants had sought out and practiced the musical stimuli from the first session. Some of the borderline-significant results might have passed the critical threshold for significance if these data could have been included. In general, there is a limitation that the lyrics were pre-existing and publicly available, and therefore no control existed regarding whether other participants had listened to the musical stimuli in between the first and second session. It is important to note that this limitation would only be relevant to data from the 1-week-delayed recall test. This concern would be completely alleviated by using original lyrics and songs, though this was addressed in the current study by omitting anyone who indicated prior knowledge of the song(s).

The results highlight several related areas that could be extended with further research. To begin, it would be beneficial to enhance understanding of the relationship between musical experience—and expertise in general—and memory tasks when performed in or outside the domain of expertise. This could also be expanded to investigate whether experts in various areas tend to spontaneously use their domains of expertise as a basis for chunking (mnemonic) strategies when learning new material. Although in the current study we did not find higher clustering scores in the music condition for musicians compared to nonmusicians, more research is needed to investigate this topic.

The motivation behind this study was to determine whether music provided effective memory support for encoding new information, using a research design that improved on prior research with regard to internal validity (e.g., a controlled laboratory setting, counterbalancing conditions, a mainly within-subjects design with repeated recall tests). As a result, this study provides stronger evidence for the advantage of music as a support for encoding new information.

A next step is to examine implications for education. Music is already being used to teach lower level educational content and is also being applied creatively to higher level content. For example, VanVoorhis (2002) demonstrated better learning and student enjoyment in a psychological statistics course when students learned statistical information through "stat jingles." Future research should investigate the advantages and boundary conditions of musical learning. With specific consideration to the fact that much of the material covered in educational settings might not fit neatly into a narrative presentation, the current study could be replicated using non-narrative material as stimuli in order to explore differences in recall or experience that may accompany recall for a set of declarations (e.g., facts, formulas) rather than a story. Recent evidence-based initiatives directed to educators have provided suggestions about the types of low-cost, high-impact learning strategies that should be implemented in the classroom (e.g., Roediger & Pyc, 2012). Educators should not minimize the importance of also choosing strategies that may feel more successful—and even enjoyable—to the learners. Learning through music may be in this category of strategies that both enhance objective measures of memory, and also enhance motivation to learn.

References

Amidzic, O., Riehle, H. J., & Elbert, T. (2006). Toward a psychophysiology of expertise: Focal magnetic gamma bursts as a signature of memory chunks and the aptitude of chess players. *Journal of Psychophysiology, 20*, 253–258. http://dx.doi.org/10.1027/0269-8803.20.4.253

Bellezza, F. S. (1996). Mnemonic methods to enhance storage and retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (pp. 345–380). http://dx.doi.org/10.1016/B978-012102570-0/50012-4

Bilalić, M., Langner, R., Erb, M., & Grodd, W. (2010). Mechanisms and neural basis of object and pattern recognition: A study with chess experts. *Journal of Experimental Psychology: General, 139*, 728–742. http://dx.doi.org/10.1037/a0020756

Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music, 31*, 123–138. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/0305735603031002290

Calvert, S. L., & Billingsley, R. L. (1998). Young children's recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology, 19*, 97–108. http://dx.doi.org/10.1016/S0193-3973(99)80030-6

Campabello, N., De Carlo, M. J., O'Neil, J., & Vacek, M. J. (2002). Music enhances learning. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED471580&site=ehost-live

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81. http://dx.doi.org/10.1016/0010-0285(73)90004-2

Chazin, S., & Neuschatz, J. S. (1990). Using a mnemonic to aid in the recall of unfamiliar information. *Perceptual and Motor Skills, 71*, 1067–1071. Retrieved from http://journals.sagepub.com/doi/pdf/10.2466/pms.1990.71.3f.1067

Cohen, J. (1988). *Statistical power analysis for the behavior sciences*. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Deason, R. G., Simmons-Stern, N. R., Frustace, B. S., Ally, B. A., & Budson, A. E. (2012). Music as a memory enhancer: Differences between healthy older adults and patients with Alzheimer's disease. *Psychomusicology: Music, Mind, and Brain, 22*, 175–179. http://dx.doi.org/10.1037/a0031118

Gingold, H., & Abravanel, E. (1987). Music as a mnemonic: The effects of good- and bad-music settings on verbatim recall of short passages by young children. *Psychomusicology: A Journal of Research in Music Cognition, 7*, 25–39. http://dx.doi.org/10.1037/h0094188

Ginsborg, J., & Sloboda, J. A. (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music, 35*, 421–440. http://dx.doi.org/10.1177/0305735607072654

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139–183. http://dx.doi.org/10.1016/S0166-4115(08)62386-9

McElhinney, M., & Annett, J. M. (1996). Pattern of efficacy of a musical mnemonic on recall of familiar words over several presentations. *Perceptual and Motor Skills, 82*, 395–400. http://dx.doi.org/10.2466/pms.1996.82.2.395

Meijs, C., Hurks, P., Rozendaal, N., & Jolles, J. (2013). Serial and subjective clustering on a verbal learning test (VLT) in children aged 5–15: The nature of subjective clustering. *Child Neuropsychology, 19*, 385–399. http://dx.doi.org/10.1080/09297049.2012.670215

Moore, K. S., Peterson, D. A., O'Shea, G., McIntosh, G. C., & Thaut, M. H. (2008). The effectiveness of music as a mnemonic device on recognition memory for people with multiple sclerosis. *Journal of Music Therapy, 45*, 307–329. http://dx.doi.org/10.1093/jmt/45.3.307

Morrongiello, B. A., & Roes, C. L. (1990). Children's memory for new songs: Integration or independent storage of words and tunes? *Journal of Experimental Child Psychology, 50*, 25–38. http://dx.doi.org/10.1016/0022-0965(90)90030-C

Racette, A., & Peretz, I. (2007). Learning lyrics: To sing or not to sing? *Memory and Cognition, 35*, 242–253. http://dx.doi.org/10.3758/BF03193445

Rainey, D. W., & Larsen, J. D. (2002). The effects of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception, 20*, 173–186. http://dx.doi.org/10.1525/mp.2002.20.2.173

Roediger III, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition, 1,* 242–248. http://dx.doi.org/10.1016/j.jarmac.2012.09.002

Scruggs, T. E., & Brigham, F. J. (1991). Utility of musical mnemonics. *Perceptual and Motor Skills, 72*, 881–882. https://doi.org/10.2466/pms.1991.72.3.1067

Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for song. *Cognition, 16*, 285–303. http://dx.doi.org/10.1016/0010-0277(84)90031-3

Silverman, M. J. (2007). The effect of paired pitch, rhythm, and speech on working memory as measured by sequential digit recall. *Journal of Music Therapy, 44*, 415–427. http://dx.doi.org/10.1093/jmt/44.4.415

VanVoorhis, C. R. W. (2002). Stat jingles: To sing or not to sing. *Teaching of Psychology, 29*, 249–250. Retrieved from http://psycnet.apa.org/psycinfo/2002-15317-012

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1471–1485. http://dx.doi.org/10.1037/0278-7393.20.6.1471

Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology, 76*, 268–275. http://dx.doi.org/10.1037/0021-9010.76.2.268

Author Note. Sarah N. Lummis, Jennifer A. McCabe, Abigail L. Sickles, Rebecca A. Byler, Sarah A. Hochberg, Sarah E. Eckart, and Corinne E. Kahler, Goucher College.

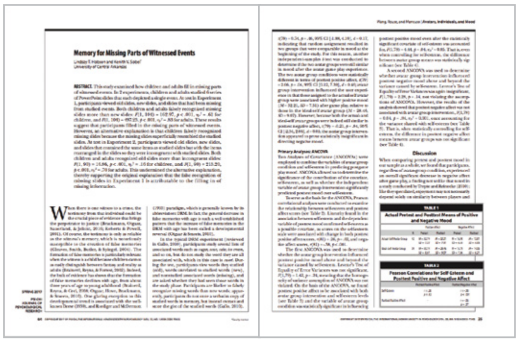
Thanks to Bianca Z. Stern, Alyssa J. Hauptman, and Talya Stern, for assistance with editing and data scoring and entry, and Brandon Meyers-Orr for assistance with designing the study. Special thanks to all *Psi Chi Journal* reviewers for their support.

Correspondence concerning this article should be addressed to Jennifer A. McCabe, Center for Psychology, Goucher College, 1021 Dulaney Valley Road, Baltimore, MD 21204, USA. E-mail: Jennifer.McCabe@goucher.edu

PSI CHI Advertising Contract: *Psi Chi Journal*

CLIENT INFORMATION	
Advertiser	
Contact Name	
Address | Street or P.O. Box	
City | State | Zip | Country	
Phone (daytime)	E-mail
Submitted by	
Authorized Signature	

DIGITAL PUBLICATION			
Issue			
❏ Spring ______	❏ Summer ______	❏ Fall ______	❏ Winter ______
Size/Dimensions			
❏ Full page (no bleed) 6⅝" x 9¼"	Cost: $400 Black & White only	❏ Half page (horizontal) 6⅝" x 4½"	Cost: $275 Black & White only
Price: $			



OUR JOURNAL

Advertising in *Psi Chi Journal* allows you to connect with established psychology researchers and mentors, as well as undergraduates and graduate students striving to build a career in one of the many areas of research. People regularly visit our journal online to

- view current and past issues,
- submit their research for publication,
- learn about reviewing for *Psi Chi Journal*,
- share invited editorials as teaching tools in the classroom.

All issues and advertisements are permanently free online to both members and nonmembers alike. During the 2015–16 fiscal year, psichi.org received **1.67 million page views**, ensuring that high-achieving students and professionals will see your content for years to come.

To further enhance the visibility of our journal, latest issues and calls for submissions are regularly featured in *Psi Chi Digest* e-mails (**180,000+ subscribers**) and on our three social media platforms.

- Facebook (22,500+ followers)
- Twitter (4,700+ followers)
- LinkedIn (10,200+ followers)

Articles are also indexed in PsycINFO, EBSCO, and Crossref databases—essential tools that researchers use to search for millions of psychology-related articles. This makes Psi Chi Journal a key place to communicate your message with our Professional Organization's **700,000+ lifetime members** and far beyond.

AD SPECIFICATIONS

Digital format: PDF, EPS, and TIFF
Resolution: 300 dpi | B&W line art—1,200 dpi Black & white ads (no RGB or 4-color process)PDF settings: Press quality, embed all fonts

DEADLINE/BILLING

Payment due upon receipt of invoice.

CONTACT

Submit contract by e-mail to
Susan Iles
Advertising Sales
Psi Chi Central Office
E-mail: psichieye@psichi.org
Phone: 423-771-9964

See past issues of *Psi Chi Journal of Psychological Research* at http://www.psichi.org/?journal_past

Stay connected with PSI CHI
www.psichi.org

All advertisements must be scholarly and professional in nature, and Psi Chi reserves the right to reject (or cancel) any ads that are not in the best interest of the Organization or consistent with the Society's mission.

ADV-JN-DG (05-2017)

PSI CHI
THE INTERNATIONAL HONOR SOCIETY IN PSYCHOLOGY

