RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 26 (2003-2004)
Indiana University

Some Effects of Early Musical Experience on Sequence Memory Spans[1]

Adam T. Tierney and David B. Pisoni

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This research was supported by NIH-NIDCD Research Grant R01 DC00111 to Indiana University. The author would like to thank David Pisoni and Tonya Bergeson for their guidance with the research and comments on earlier versions of this paper.

Some Effects of Early Musical Experience on Sequence Memory Spans

Abstract. Language and music were once considered completely separate systems, located in the left and right hemispheres of the brain, respectively. However, recent evidence from neuroimaging and behavioral studies has suggested that music and language may be closely linked. For example, syntactic processing and the processing of harmony may share certain neural resources, and some aspects of speech and music may be processed by overlapping working memory systems. If such overlap does exist, early experience with music and musical performance may affect not only the cognitive skills used to process music, but also language skills. The present study tested experienced musicians and three groups of musically inexperienced subjects, gymnasts, Psychology 101 students, and video game players on digit span, word span and sequence memory span tasks. By including skilled gymnasts who began studying their craft by age six, video game players, and Psychology 101 students, we attempted to control for some of the ways skilled musicians may differ from the general population, in terms of gross motor skills and intensive experience in a highly skilled domain form an early age. We found that musicians displayed higher memory spans than the comparison groups on digit span and auditory conditions of the sequence span task, but no differences were found between the four groups on the visual condition of the sequence span task. These results provide further converging support to recent findings showing that musical experience and activity may affect verbal rehearsal, phonological coding, and the allocation of attention in sequence memory tasks.

Introduction

The idea that exposure to or training in music can make us smarter is perhaps most often associated in the public mind with the so-called "Mozart Effect." This effect was first reported by Rauscher, Shaw and Ky (1993) who found that exposure to music written by Mozart, as compared to exposure to silence, led to increased scores on the Paper-Folding and Cutting subset of the Stanford-Binet IQ test (Thorndike, Hagen, & Sattler, 1986). Follow-up studies, however, have had a great deal of difficulty replicating the effect. Chabris (1999), in a review of 20 studies of the Mozart Effect, found an average cognitive enhancement of only 1.4 IQ points, while Thompson, Schellenberg, and Husain (2001) found that the cognitive enhancement was not reliable when arousal scores, enjoyment ratings, or subjective mood-arousal ratings were controlled. Nevertheless, the claim that musical experience can lead to cognitive facilitation lives on under a number of different guises. In particular, some researchers, pointing to evidence for a stronger neural link than previously supposed between music and language, argue that musical training facilitates verbal skills, especially verbal memory (Chan, Ho, & Cheung, 1998; Ho, Cheung, & Chan, 2003; Jakobson, Cuddy, & Kilgour, 2003; Koelsch et al., 2002; Patel et al., 1998).

Until recently, one of the only techniques available for studying possible neural overlap between music and language was to examine clinical cases of amusia and aphasia. Marin and Perry (1999) provide a comprehensive overview of these case studies, including several in which the patients suffered from amusia without aphasia, and vice versa. Other cases cited by Marin and Perry include disrupted memory for melodies and disorders of rhythm production with no aphasia, verbal alexia accompanied by preserved music reading and writing, and auditory agnosia restricted to speech sounds that did not impair nonverbal discrimination.

Additional evidence against overlap of music and language processing comes from a substantial body of research showing that music processing takes place primarily in the right hemisphere of the brain, while speech processing takes place primarily in the left. Zatorre and Sampson (1991), for example, asked patients with damage to the right or left anterior temporal lobes to compare two tones separated by six distractor tones, then to indicate if the tones were the same or different. Only the patients with right hemisphere damage showed inhibition on this task. In another study, Penhue, Zatorre, and Evans (1998) had subjects reproduce simple rhythms by tapping and found that the right planum temporale, a part of the auditory association cortex, experienced the greatest activation. Finally, Tervaniemi et al. (2000) used PET to investigate how subjects process musical sequences of “A” major chords interspersed with occasional “A” minor chords, as well as linguistic sequences of standard and occasional deviant phonemes. Their results showed that the deviant chords activated the right superior temporal gyrus most strongly while the deviant phonemes activated the left superior temporal gyrus most strongly.

These studies all support the proposal that music and language are processed by two independent systems, each modality-specific. However, the three studies cited all suffer from a common design problem: they used only non-musicians as subjects. Studies using skilled musicians as subjects have shown greater left hemisphere activation during tonal processing tasks, especially if the task is difficult. For example, Messerli, Pegna, and Sordet (1995) presented musicians and non-musicians with the melodies from eight popular folk songs and then asked both groups of subjects to point to one of eight pictures that best matched the song. Musicians displayed a right ear advantage while non-musicians displayed a left ear advantage. Also, Harris and Silberstein (1999) asked female musicians to memorize ten-note patterns while using EEG to measure their brain activity. The subjects were then presented with a part of the pattern and were asked to determine if any of the notes had been changed. The authors found that musicians showed mainly left temporal activation during this task and the effect was positively correlated with the difficulty of the task and subjects’ degree of musical experience.

Thus, research involving skilled musicians has found that music processing is not confined to the right hemisphere, a finding that breathed new life into the question of possible neurobiological linkage between language and music processing. One of the first studies to set foot in this new area of research was reported in Patel, Gibson, Ratner, Besson, and Holcomb (1998). They found that the P600, an ERP correlate of processing of syntactic incongruities, is not specific to language but can also be elicited by chordal sequences containing harmonic incongruities. The amplitude of the P600s was correlated with the degree of violations of expectations in the sentences and chord sequences. Follow-up studies by Maess, Koelsch, Gunter, and Friederici (2001) and Koelsch et al. (2002) using MEG and fMRI respectively, found that processing of harmonic incongruities was also associated with activation in Broca’s Area, Wernicke’s Area, and the superior temporal sulcus.

While some researchers have investigated whether complex musical and linguistic abilities are processed in part by overlapping neural systems, other researchers have assessed whether musical and linguistic information are stored in a single auditory system within working memory or in two separate subsystems. In an early study, Deutsch (1970) played eight tones for subjects, then asked them to compare the first and last tone to determine if they were the same or different. In some conditions, however, spoken numbers replaced the middle six tones. The interpolated tones greatly impaired subjects’ performance, but the spoken numbers had no significant effect on performance. Deutsch concluded that a separate system exists for the storage of tonal pitch in working memory.

Pechmann and Mohr (1992), however, contested Deutsch’s conclusions. The authors used a similar experimental paradigm, but presented interpolated tones, spoken numbers, or visual stimuli between the two tones. Subjects included both instrumental musicians and non-musicians. While only the interpolated tones inhibited the musicians’ performance, the tones, spoken numbers, and visual stimuli all significantly inhibited non-musicians' performance. The authors speculated that because non-musicians were less efficient in processing the tones, they were more susceptible to interference, even from less related stimuli.

Other researchers have attempted to determine the effects of hearing instrumental and vocal background music on various verbal memory tasks. Salamé and Baddeley (1989), for example, tested 24 non-musicians on recall of sets of nine visually presented digits. While the subjects performed this task, they heard instrumental background music, vocal background music, or silence. The subjects' performance was significantly inhibited by the presentation of instrumental music, but vocal music caused an even greater degree of interference.

However, this research on instrumental music and task interference ignores the fact that music consists of a number of different components. For instance, instrumental music contains significant rhythmic information—if music is stored in a "tonal loop," does it follow that rhythm is also stored there? Saito and Ishio (1998) addressed the question of whether rhythmic information is stored separately in working memory from phonological information. The researchers asked non-musicians to reproduce rhythmic patterns by pressing a key on a computer. During presentation of the rhythmic patterns, subjects were asked to either mouth the vowels "A-E-I-O-U" repeatedly, draw squares as many times as possible, or focus their attention on the rhythms. Performance in the concurrent articulation condition was worse than in the spatial task condition, which was worse than the control condition. The authors concluded that rhythmic information is retained in working memory via the phonological loop.

Taken together, the studies with musicians indicate that music and language may be processed by partially overlapping systems in the brain, both at the level of simple abilities, such as maintaining auditory information in working memory, and at the level of more complex abilities such as syntax and harmony processing. Moreover, research has shown that early experience with music can have effects on brain development that are strong enough to be detectable in the physical brain structure. In 1995, for example, Schlaug, Jancke, Huang, and Steinmetz, using MRI, found that a group of musicians with perfect pitch had a greater disparity between the left and right planum temporale (PT), as compared to non-musicians and musicians with relative pitch. The authors speculated that, given that PET has demonstrated that the PT is part of an area that is involved in music perception, this increase in left-right disparity might indicate increased lateralization of music processing in musicians with absolute pitch, as compared to non-musicians and musicians with relative pitch.

In a follow-up study, Schlaug, Jancke, Huang, Staiger, and Steinmetz (1995b) reported that the midsagittal area of the anterior half of the corpus callosum was larger in a group of musicians who had begun playing piano and stringed instruments by age seven than in a group of non-musicians. Using fMRI, Thomas, Pantev, Wienbruch, Rockstroh, and Taub (1995) found that string musicians had increased cortical representation of the fingers of their left hands, as compared to non-musicians. This finding was correlated with the age at which the person began to play music. Finally, Schneider et al. (2002), using fMRI, found that musicians, as compared to non-musicians, had a significantly greater volume of gray matter in Heschl's gyrus, especially in the anteromedial portion (that is, the primary auditory cortex).

Given the findings that early experience with music can cause measurable neurobiological changes, and the results that music and language processing may be psychologically and neurally linked, it is reasonable to ask what differences in verbal skills musicians might show when compared to non-musicians. Most prior research in this area, although suggestive, did not include appropriate comparison groups, but simply compared subjects with musical experience to subjects without musical experience. Any differences found in verbal skills could therefore be confounded by differences in any structured domain obtained from the early age at which most musicians begin playing. Confounding variables could include socioeconomic status, innate motor skills, or training effects.

The first researcher to attempt to correlate musical and non-musical cognitive skills was McMahon (1982). He found a positive correlation between performance on the Enticknap Picture Vocabulary Test, which requires children to name line drawings of objects on 48 test cards, and a tonal discrimination test designed by the experimenter. Further correlations between verbal and musical skills were discovered by Atterbury (1985), who reported that learning-disabled readers performed significantly worse on tonal discrimination and rhythmic discrimination tasks. Barwick, Valentine, West, and Wilding (1989) also found significant positive correlations between a tonal memory test and reading age, even after adjusting for chronological age and IQ. Finally, Lamb and Gregory (1993) reported significant positive correlations between a pitch discrimination test and tests requiring children to read nonsense words and demonstrate an understanding of rhyme and alliteration.

These correlational studies suggest that basic musical abilities, such as tone, rhythm, and chord discrimination, are related somehow to reading ability. However, the children with better musical skills may also possess better language skills because children with higher socioeconomic status and a more intellectual home environment may be exposed more often to certain aspects of language and music (their parents may read to them more often, for example). In an effort to move beyond mere correlation and determine whether or not early musical training actually affects subjects' non-musical cognitive skills, several researchers have turned to quasi-experimental studies in which subjects with musical experience are directly compared to subjects without musical experience.

In the first study of this kind, Huntsinger and Jose (1991) found that musically experienced children performed better than musically inexperienced children on tonal memory and digit span tasks. Similarly, Chan, Ho, and Cheung (1998) reported that musicians recalled more words from a visually presented 16-word list than did non-musicians. Kilgour, Jakobson, and Cuddy (2000) asked undergraduate musicians and non-musicians to memorize lyrics that were either spoken or sung, one line at a time. The musicians showed better recall of both the sung and spoken lyrics. Munzer, Berti, and Pechmann (2002) found that musicians performed better than non-musicians on a task which required them to decide if the first and last stimuli in a set of speech sounds were the same or different.

Finally, Jakobson, Cuddy, and Kilgour (2003) attempted to establish a theoretical basis for the finding that music training can lead to enhanced verbal recall. The researchers compared musicians and non-musicians on tests of verbal recall and temporal order processing, in which the subjects were asked to determine the order of presentation of two tones or two syllables of equal duration. The correlation between years of musical training and verbal recall scores was strong and significant, as was the correlation between musical training and the composite score for the temporal order tests. Using a multiple-regression analysis, the authors found that the relation between verbal recall scores and years of music training was reduced when temporal-order processing was included in the regression equation. Thus, it appears that the advantage that musicians show in verbal recall tasks may be due in part to their enhanced temporal order processing abilities.

However, one cannot infer causation from these results because it is possible that children with greater cognitive skills in the verbal/auditory domain are simply more likely to be drawn to music and that this predilection could account for the musicians' better performance on verbal memory tasks. In an attempt to establish whether or not the better performance on reading and verbal memory tasks that has been previously found actually results from musical training, several other researchers have performed more extensive, truly experimental studies with randomly selected experimental and control groups. Ho, Cheung, and Chan (2003), for example, found that a year of piano instruction led to improved verbal recall scores, as compared to a group that received no instruction. More recently, Schellenberg (2004) provided children with a year of piano instruction, a year of drama lessons, or no instruction, and found that the children who received piano instruction showed significantly greater increases in IQ than the control groups; this difference was found on all but two of the 12 subtests of the WISC-III IQ test (Wechsler 1991).

It is difficult to say whether these differences in verbal skills are truly due to musical training, because skilled musicians differ from non-musicians in many other ways unrelated to music, such as motor skill and socioeconomic status. One way to attempt to account for some of these possible differences between musicians and non-musicians is to test the verbal skills of a comparison group that has practiced a skilled motor task intensively from an early age. For this reason, we compared the ability of skilled musicians from the Indiana University School of Music to perform a number of short-term memory and working-memory tasks in both the spatial and verbal/auditory domains with that of subjects from three comparison groups, all of whom were unable to read music: (a) gymnasts from the IU Gymnastics Club, (b) Psychology 101 students at IU, and (c) students who played video games for at least 1.5 hours per day. Green and Bavelier (2003) found that video-game players showed facilitated processing on tests of visual attention and rapid temporal processing. It is possible, therefore, that video-game players would show differential facilitation to visual and auditory patterns. These four groups were tested on a number of short-term memory and working-memory tasks in both the visual-spatial and verbal-auditory domains to determine what effect, if any, prior early musical training and experience had on their ability to encode and recall information from immediate memory.

Methods

Subjects

Forty-five students participated in the study. A summary of the demographic characteristics of the participants is given in Table 1. By sending e-mails to the distribution list of the IU Gymnastics Club, we recruited ten gymnasts whom could not read music and started gymnastics when they were 4.5 years of age. They practiced an average of 4.5 hours per week. We also recruited 12 experienced music students who were piano performance majors by sending e-mails to the distribution list for all of the undergraduate and graduate piano performance majors. The mean age at which the pianists began playing the piano was 7 years. The average number of hours they spent practicing every week was 24.18. Twelve students were unselected undergraduates who could not read music and were enrolled in beginning

	Age	SAT Verbal	SAT Math	Num. Males	Num. Females
Gymnasts	19.9 (SD 1.4)	536 (SD 36)	533 (SD 39)	0	10
Musicians	21.1 (SD 2.3)	650 (SD 51)	638 (SD 75)	7	5
Non-musicians	19.4 (SD 1.2)	530 (SD 105)	540 (SD 96)	2	10
Video gamers	22.3 (SD 3.1)	540 (SD 85)	576 (SD 85)	11	0

Table 1. Demographic data for gymnasts, musicians, Psychology non-musicians and video game players.

psychology courses at IU. The psychology students were contacted via a list of students who indicated on a questionnaire at the end of class that they wished to participate in paid psychology experiments. Finally, we recruited 11 students who played video games for at least 1.5 hours a day, and were unable to read music. The difference in the ages between the four groups at the time of testing was not statistically significant ($p > .05$). All subjects were asked to report their Math and Verbal SAT scores, so we could assess any differences in academic and intellectual abilities between the groups. All subjects were paid $15 for their time.

Stimulus Materials

The subjects completed several short-term memory and working memory tasks including a digit span task, a word span task, and a word familiarity test. For the digit span task, ten spoken digits were acquired from the Texas Instruments 46-Word (TI46) Speaker-Dependent Isolated Word Corpus (Texas Instruments, 1991). Test words for the word span tasks were obtained from a digital speech database (Goh & Pisoni, 2003) containing the 300 monosyllabic English words from the Modified Rhyme Test (House, Williams, Hecker & Kryter, 1965) and from phonetically balanced word lists. Sixty-six "easy" consonant-vowel-consonant (CVC) words and 66 "hard" CVC words, whose lexical difficulty was computed according to the neighborhood activation model (Luce & Pisoni, 1998), recorded by a single male voice were chosen so that the two sets consisted of different neighborhood density and neighborhood frequency but did not differ on word frequency and the number of intra-set neighbors.

The test words for the word-familiarity (FAM) vocabulary test were taken from Stallings, Kirk, Chin, and Gao (2002), a shortened version of the FAM test originally developed by Lewellen, Goldinger, Pisoni, and Greene (1993). The 150-word stimulus list consisted of 50 high-familiarity words, 50 medium-familiarity words, and 50 low-familiarity words. The familiarity scores were based on normed familiarity ratings of 20,000 words collected originally by Nusbaum, Pisoni, and Davis (1984).

Experimental Design and Procedures

Digit Span Task. Subjects were asked to recall lists of spoken digits, presented over headphones. The list length was increased by one item after every two trials, starting at a list length of four and ending with a list length of ten, with a total of 14 trials for each test. For the first test, Forward Digit Span, subjects were asked to recall and write down the digits as they were given. For the second test, Backward Digit Span, subjects were asked to recall and write down the digits in the reverse order.

Word Span Task. Subjects were asked to recall lists of spoken words, presented over headphones. The list length was once again increased by one item after every two trials, starting at a list length of three words and ending with a list length of eight, for a total of 12 trials for each test. Half of the subjects from each group completed the "easy" word span test first and the other half completed the "hard" word span test first.

Word Familiarity Ratings. This is a test of subjective familiarity with easy, medium difficulty, and hard words. Fifty words from each category were presented randomly on a computer screen, one at a time, for a total of 150 words. The subjects were asked to judge their familiarity with the word by pressing the appropriate number key on the keyboard, from "1" for least familiar to "7" for most familiar.

Sequence Reproduction Span Task. This task used a modified version of the Simon Memory Game, a round box with four colored panels (Karpicke & Pisoni, in press; Pisoni & Cleary, 2004). In the first condition, visual-only (VO), the colored panels lit up, one at a time, in a sequence. The subjects were then asked to reproduce the sequence of colored lights by pressing the appropriate panels on the box. In the second condition, auditory-only (AO), the subjects heard a sequence of color names over their headphones. They were asked to reproduce the sequence by pressing the appropriate panels on the box. In the final condition, audiovisual (AV), the panels on the box were illuminated and the subjects simultaneously heard names of colors that corresponded to the colors of the lights. In each of the three presentation conditions, the sequence length started out at one item, then increased by one whenever the subjects reproduced a particular sequence length correctly twice in a row. If the subjects reproduced a sequence incorrectly, the sequence length decreased by one item. Each condition (VO, AO, and AV) consisted of twenty trials.

All subjects first completed the forward and backward digit span tasks, followed by the easy and hard word span tasks. After a five-minute break, the subjects then completed the word familiarity task and the Simon sequence memory task.

Results

Demographic Data

A series of one-way ANOVAs were performed on the subject demographic data. SAT scores were reported by 67% of musicians, 72% of gymnasts, 80% of Psychology students, and 63% of video game players. Mean SAT verbal scores were found to be significantly higher for the musicians than for the gymnasts, Psychology students, and video game players (p = .003, .002, and .003, respectively.) Mean SAT math scores were significantly higher for the musicians as compared to the gymnasts (p = .014).

Forward and Backward Digit Spans

Due to a corrupted file, the digit span data for one of the gymnast subjects was lost; therefore, the N for the gymnasts on the digit span tasks was 9 not 10. Four scores were computed to measure digit spans. "High" was the longest span which the subjects reached without making any errors. After making an error, 0.5 was added for every full list repeated correctly, resulting in the "Strict" score. For the "Absolute" score, the number of items in each list recalled correctly was summed, and finally, for the "Total" score, the total number of items recalled correctly was computed. Only the results of the "Total" scores will be reported here because the other scores did not vary widely enough to show consistent effects. Figure 1 shows mean scores for the four subject groups in both the forward and backward digit span recall conditions.

The two digit span scores were entered into a repeated measures ANOVA, with group as a between-subjects variable and type of span (forward versus backward) as a within-subjects variable. Main effects of span [$F(1,40) = 61.73$, $MS_e = 3205.004$, $p < .001$] and group [$F(3, 40) = 4.994$, $MS_e = 935.600$, $p = .005$] were found. Post-hoc Bonferroni tests revealed that musicians significantly outperformed the gymnasts on the forward digit span (p = .011) and significantly outperformed the gymnasts and the P101 non-musicians on the backward digit span (p = .049 and .014, respectively). No other significant differences in performance were found between the video game players and the remaining three groups for either forward or backward span.

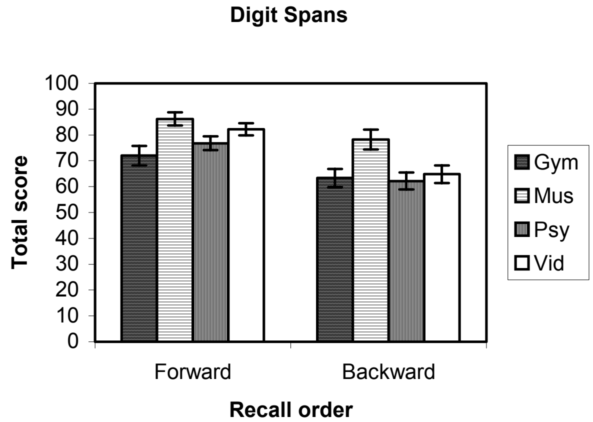

Figure 1. Mean total span scores for the four groups in both the forward and backward digit span recall conditions. Error bars are standard error.

Word Span

Scoring for the word span task was identical to the procedure used in the digit span task. Figure 2 shows mean "Total" scores for the four subject groups in both of the conditions (easy and hard word span). The word span scores were entered into a repeated measures ANOVA, with group as a between-subjects variable and type of span (hard vs. easy) as a within-subjects variable. A significant interaction between span and group was found [$F(3,40) = 3.072$, $MS_e = 87.517$, $p = .039$.] No main effects of group or span type were found. Post-hoc analysis revealed that musicians performed significantly better than the video game players on the hard word span condition ($p = .014$); none of the other differences reached significance.

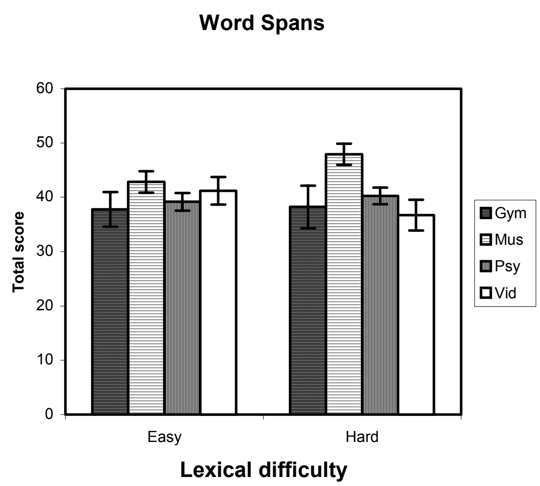

Figure 2. Total items recalled correctly for the four groups of subjects for the two types of word span. Error bars are standard error.

Word Familiarity Ratings

Figure 3 shows mean familiarity ratings for the three subject groups in each of the three different word familiarity conditions (high-familiarity, medium-familiarity, and low-familiarity). The FAM ratings were entered into a repeated measures ANOVA, with group as a between-subjects variable and word familiarity (low, medium, and high) as a within-subjects variable. A main effect of familiarity [$F(2,60) = 767.037$, $MS_e = 168.290$, $p < .001$] was found. Post-hoc analysis revealed no significant differences between the four groups. Musicians were not more likely to rate words of varying difficulty as more familiar than the three comparison groups.

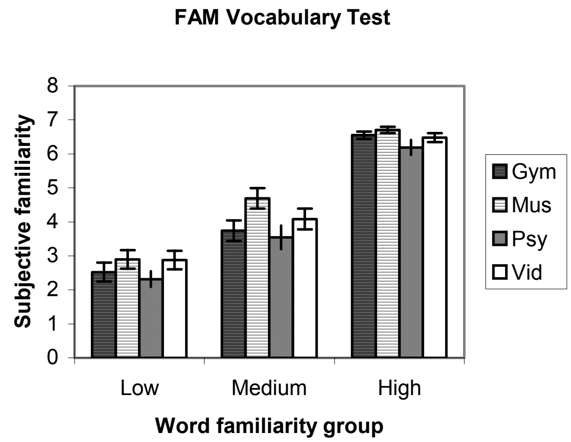

Figure 3. Mean subjective familiarity ratings for the four groups in the FAM rating task. Error bars are standard error.

Simon Sequence Memory Span

Four different memory span scores were computed: ALL was the longest sequence which the subjects reproduced correctly 100% of the time, HALF was the longest span which the subjects reproduced correctly 50% of the time, ONCE was the longest span which the subjects reproduced correctly at least once, and WEIGHT was a weighted sum of the percentage of correct responses at each list length (see Pisoni & Cleary, 2004). Analyses were carried out on each of the four different span methods. A similar pattern of results was found for each type of score, so only the results of the weighted score will be reported here. Figure 4 shows mean weighted span scores for the four subject groups in each of the three different presentation conditions (VO, AO, and AV).

The Simon weighted span scores were entered into a repeated measures ANOVA, with group as a between-subjects variable and presentation condition (VO, AO, and AV) as a within-subjects variable. Main effects of Simon condition [$F(2,82) = 16.88$, $MS_e = 4.307$, $p < .001$] and group [$F(3, 41) = 4.565$, $MS_e = 4.548$, $p = .008$] were found. A significant interaction between condition and group [$F(6, 82) = 3.803$, $MS_e = .970$, $p = .002$] was also found. None of the pairwise differences in VO were significantly different. For the AO condition, the musicians (mean score 6.48, *SD* =.92) outperformed the gymnasts (mean score 5.43, *SD* =.63, $p = 0.006$), the psychology students (mean score 5.02, *SD* =.33, $p < 0.001$), and the video game players (mean score 5.29, SD =.65, $p = .001$). No significant difference in performance was found between groups for the AV condition, although there was a numerical trend for the musicians to display longer spans in this condition too.

Post hoc Bonferroni tests revealed that, overall, subjects did not perform better on the AO condition than on the VO condition ($p = .11$). However, they did show significantly better performance on the AV condition as compared to the VO condition ($p < .001$) and the AO condition ($p < .001$). Further analysis revealed that the difference between the AO and AV conditions was only significant for the psychology subjects ($p = .002$). The difference between the VO and AO conditions was significant only for the musicians (p = .011). Thus, among the four groups of subjects only the musicians were able to reproduce auditory patterns better than visual patterns.

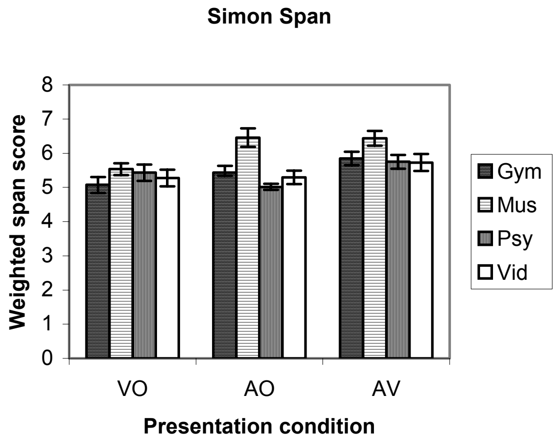

Figure 4. Weighted span scores for the four groups of subjects in the three different conditions (VO, AO, and AV) of the Simon Span task. Error bars are standard error.

Correlations

A summary of the intercorrelations between the four different tests can be found in Table 2. Simon VO, AO, and AV conditions were all significantly correlated with one another. The AO and AV conditions also correlated significantly with the Backward Digit Span measure, but not with the Easy and Hard Word Span measures or the Forward Digit Span measure. The Digit Span and Word Span tasks were all significantly correlated with one another.

	Digit Span		Word Span		FAM Familiarity Test			Simon Span		
	FDS	BDS	Easy	Hard	FAM-low	FAM-mid	FAM-high	VO	AO	AV
Forward Digit Span	1									
Backward Digit Span	.634**	1								
Easy Word Span	.630**	.428**	1							
Hard Word Span	.432**	.483**	.557**	1						
Familiarity test--low	.137	.193	.002	.131	1					
Familiarity test--mid	.256	.362*	.045	.182	.868**	1				
Familiarity test--high	.165	.274	.060	.063	.587**	.809**	1			
Simon VO	.213	.236	.213	.292	-.151	-.147	-.148	1		
Simon AO	.292	.428**	.054	.134	.062	.183	.173	.352*	1	
Simon AV	.274	.351*	.205	.192	-.179	-.041	-.048	.505**	.665**	1

Table 2. Pearson correlations. A (*) signifies that the correlation is significant at the .05 level; (**) signifies that the correlation is significant at the .01 level.

Discussion

Skilled musicians performed significantly better than the gymnasts, introductory psychology students, and video game players on the auditory-only (AO) condition of the Simon memory span. No significant differences were found on the visual-only (VO) condition or the audio-visual (AV) condition of this task. These new findings on sequence memory span suggest that skilled musicians display a greater capacity for reproducing randomized auditory sequences than non-musicians. The present findings on sequence memory span provide an excellent counterargument to those who would claim that any cognitive differences between skilled musicians and non-musicians might simply be due to general intelligence, or to global differences in academic achievement. Effects of general intelligence and education are not likely to be limited to a single domain or sensory modality (Gardner 1993). Moreover, the facilitation of musician performance in the AO condition cannot be due to differences in motor skills, because the output for each of the three conditions involved exactly the same response demands: pressing a sequence of buttons on the Simon box. The only difference between the three conditions was the modality of the input signals.

The data obtained from the skilled musicians for the three Simon conditions also showed patterns that were different from those found in previous Simon experiments that used psychology students from the general university population. Pisoni and Cleary (2004) reviewed several studies that tested normal-hearing adults and normal-hearing 8- and 9-year-old children on the AO, VO, and AV conditions of the Simon memory span. Children performed worse overall than adults, but both groups showed two significant effects. First, both groups of subjects displayed a "modality effect: performance was better on AO than on VO. Secondly, both groups of subjects exhibited a "redundancy gain": performance was better in the AV condition than in either the VO or AO conditions. Both findings suggest that sequence memory spans are not only sensitive to input modality but also display cross-modal interactions between auditory and visual signals.

In contrast to Pisoni and Cleary's earlier findings with adults and normal-hearing children, in the present study skilled musicians showed a strong redundancy gain when the AV condition was compared to the VO condition, but no redundancy gain when the AV condition was compared to the AO condition. Of the four groups tested in this experiment, only the skilled musicians exhibited a significant modality effect, with performance on AO greater than performance on VO, although the gymnasts showed a trend in the same direction. When performance was averaged across all four groups, we did find an overall redundancy gain (that is, performance on AV was better than on VO and AO), but we found no significant modality effect. These differences may be due to the small number of subjects that we were able to recruit for this study. Pisoni and Cleary (2004) did find a strong modality effect, but they used 48 subjects in the study.

Although we replicated the redundancy gains reported earlier by Pisoni and Cleary (2004) and found a trend towards a modality effect, skilled musicians displayed little benefit from the redundant visual information presented in the AV condition as compared with the AO condition. The other three groups of subjects all showed facilitation of performance in the AV condition when compared to the AO condition, though this trend reached significance only for the psychology subjects. The auditory information was so salient to the musicians that adding redundant visual information provided little additional benefit. Alternatively, the musicians may have used a different encoding strategy in the AO condition than the other groups, which relied more heavily on the temporal coding of auditory information in these sequential patterns.

It is possible that the lack of a redundancy gain for musicians when AO was compared to AV was due to a tendency of musicians to selectively focus their attention on auditory patterns, even random auditory patterns, to the exclusion of other sources of sensory information. Future studies could test this hypothesis directly by decoupling the auditory and visual input in two "A–V" selective attention conditions. In one condition, the subjects would be asked to reproduce the visual pattern (i.e., VO) and selectively ignore irrelevant auditory input (that is, the names of colors that are not congruent with the colored lights). In another condition, the subjects would be asked to reproduce the auditory patterns (i.e., AO) and selectively ignore incongruent visual input. If musicians' attention is controlled or modulated by the auditory input, their performance should be facilitated in the A-V condition in which they are asked to reproduce the auditory input and ignore the visual input, but inhibited in the V-A condition in which they are asked to reproduce the visual input and selectively ignore the auditory input. In a recent study on automaticity in skilled musicians, Stewart, Walsh, and Frith (2004) asked subjects to read numbers from 1 to 5 embedded in five different musical notes; the numbers indicated which finger subjects should use to press a key. Skilled pianists, but not musically inexperienced subjects, were facilitated when the irrelevant musical stimuli were congruent with the numbers, but showed interference when the musical and numerical stimuli were not congruent. These results demonstrate that musical experience can affect the automaticity of specific information processing tasks; just as musicians in this study were not able to ignore the irrelevant musical notation, it is possible that musicians in an A–V Simon span condition might not be able to ignore irrelevant auditory information.

The musicians' long experience and activities in listening to and actively playing music appears to facilitate verbal auditory processing of temporal sequences. It is not clear, however, whether this effect is due to enhanced encoding strategies during early perceptual analysis or are the result of differences in the musicians' use of the central executive, that is, to enhanced manipulation of auditory-verbal information actively maintained in short-term memory. The differences in performance found between the groups could also be due to storage or processing activities. The digit span results can be used to try to further pinpoint the exact locus of the musicians' verbal memory facilitation. Musicians outperformed the gymnasts on the forward digit span. The advantage shown by musicians was slightly greater for the backward digit span task: musicians performed significantly better than both the gymnasts and the psychology subjects. There was, however, no interaction between span and group.

Because there was no interaction between span and group in the digit span task, it is possible that the musicians' improved performance on the Simon AO and digit span tasks may simply be due to more efficient phonological encoding, rather than active manipulation of information in working memory. Data from neuro-imaging studies have suggested that backward digit span is a measure of working memory span rather than passive short-term memory. Several recent fMRI studies have revealed that storage of information is associated with increases in regional cerebral blood flow (rCBF) in ventrolateral frontal cortex, which is found in both short-term memory tasks and working memory tasks. Manipulation and active processing of information, on the other hand, is associated with increases of rCBF in dorsolateral frontal cortex, which is generally found only in tasks thought to require the use of the central executive to actively manipulate information maintained in short-term memory (Owen, 2000). Researchers have found that the backward digit span tasks lead to activation of both ventrolateral and dorsolateral prefrontal cortex, while the forward digit span leads to activation of ventrolateral prefrontal cortex only (Owen, Lee, & Williams, 2000). Factor analysis performed on studies examining subjects' performance on various tasks thought to tap into working memory and short-term memory has provided support for the dissociation between these two memory processes, and has revealed that working memory more often correlates with measures of reading comprehension such as Verbal SAT scores (LaPointe & Engle, 1990; Cantor, Engle, & Hamilton, 1991; Kail & Hall, 2001).

The results of this study suggest that musicians show enhanced encoding of information, rather than enhanced ability to actively manipulate information. It is still not clear, however, whether the Simon sequence memory span task, with which the largest effects of musical experience are found, is solely a test of encoding or if it also involves a significant processing and manipulation component. The AO condition, for example, requires subjects not only to maintain the list of color names in short-term memory but also to convert the color names into the corresponding locations on the button box and execute the motor pattern necessary to activate the buttons. Correlational analysis on the results of this study revealed that scores on both the AO and AV conditions correlated significantly with the backward digit span task, but not with the forward digit span task.

Several differences in performance were found between the four groups on the word span task. Musicians performed significantly better than the video game players on the hard word span task, and a trend was found for better performance by the musicians than the gymnasts on the word span task. However, no significant effects were found on the easy word span task. The difficulty in finding any differences between musicians and non-musicians on the words span tasks may be due to the differences in performance between open-set and closed-set tasks (Sommers, Kirk & Pisoni, 1997); it is possible that the musicians were facilitated only on matching of auditory stimuli to stored patterns, and that when the lexicon must be accessed this advantage vanishes. On the other hand, the difficulty in finding effects may be due to the small number of subjects, and increasing the power of the study might bring the difference between musicians and psychology subjects to significance. Subjects did not show a decrement in performance on the “hard” word span task as compared to the “easy” word span task, a somewhat anomalous result. Goh and Pisoni’s (2003) study, which did find a decrement in performance for the “hard” words as compared to the “easy” words, used 56 subjects and counterbalanced the conditions. The current study, however, used 45 subjects and always presented subjects with the “easy” condition first, followed by the “hard” condition. A practice effect may, therefore, have led subjects to perform better on the “hard” condition than they would have in an experiment with counterbalanced conditions.

The results obtained from the comparison group of students who began practicing gymnastics from an early age suggest that the differences between musicians and non-musicians on the span tasks are not the result of gross motor skills or knowledge of a skilled domain from an early age. Future examinations of the effects of early music experience would do well to include multiple comparison groups drawn from other specialized subsets of the population. A 1997 survey of 663 children taking private piano lessons, for example, revealed that the children’s parents were mainly college-educated, professional, upper to upper-middle income, Caucasian suburbanites (Duke, Flowers, and Wolfe, 1997). If proposed studies of the cognitive differences between musicians and non-musicians do not attempt to take into account the many extramusical ways in which musicians differ from the general population, it is difficult if not impossible to determine precisely whether or not the results are in fact due to the subjects’ selective experience and activity with music.

Several important differences between the skilled musicians and the three comparison groups remain, and the number of subjects was too small to adjust for even those differences that could be measured. The most prominent difference between the subject groups was found in their SAT scores. Musicians had significantly higher SAT scores than the gymnasts and psychology students, both for the Math and Verbal sections. (The effect was larger for the Verbal scores than it was for the Math scores.) This result offers the greatest challenge to the claim that the Simon AO results reflect a difference that stems primarily from the intensive experience that the musicians have had with auditory stimuli from an early age. One could argue, then, that the musicians have been exposed to better quality education and, perhaps, a more intellectual home environment, that this difference has increased the amount of time they have spent on homework and reading, and that this experience with processing visually presented verbal stimuli (that is, reading) has somehow fine-tuned their ear for speech. The differences in SAT scores may also reflect the very selective nature of the School of Music at Indiana University, which is more competitive in admission than the College of Arts and Sciences. According to this argument, the differences in auditory performance between the musicians and non-musicians are not the result of musical training or selective experience with sound and auditory patterns but instead stem from the musicians' higher socioeconomic status and greater educational opportunities.

This hypothesis, however, does not account for the selective failure to find any difference at all between the four groups on the VO condition. If overall educational achievement is responsible for the musicians' enhanced auditory memory spans, then their visual memory spans in VO should have been improved as well. Many academic fields, especially mathematics, involve the use of visual-spatial processing; therefore, the difference in quantitative SAT scores should go hand-in-hand with differences in visual working memory capacity. The argument that these differences are due to educational history could be addressed in a follow-up study with psychology students whose SAT scores closely matched to those of the musicians. The group of video game players was included in an attempt to find subjects who would display precisely this pattern of performance, but they did not perform significantly better on the VO condition than the other three groups of subjects.

In conclusion, skilled musicians displayed significantly longer auditory reproductive sequence memory spans than three groups of non-musicians: gymnasts, introductory psychology students, and video game players, all who were unable to read music. This effect was highly selective in nature and was observed only for the AO condition, and not for the VO and AV conditions. Musicians also performed significantly better on Forward and Backward Digit Span tasks than the other three comparison groups. Analyses of the digit span measures revealed no significant interaction between span and group. Thus, it appears that early experience with music facilitates performance on memory span tasks and because it affects both forward and backward digit spans, it is therefore more likely to be related to early encoding and verbal rehearsal-maintenance rather than active manipulation or transformation of pattern information already in short-term memory.

References

Atterbury, B. (1985). Musical differences in learning-disabled and normal-achieving readers, aged seven, eight, and nine. *Psychology of Music*, *13*, 114-123.

Barwick, J., Valentine, E., West, R., & Wilding, J. (1989). Relations between reading and musical abilities. *British Journal of Educational Psychology*, *59*, 253-257.

Cantor, J., Engle, R.W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, *15*, 229-246.

Chabris, C.F. (1999). Prelude or requiem for the 'Mozart effect'? *Nature*, *400,* 826-827.

Chan, S., Ho, Y., & Cheung, M. (1998). Music training improves verbal memory. *Nature*, *396*, 128.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, *168*, 1604-1605.

Duke, R., Flowers, P., & Wolfe, D. (1997). Children who study piano with excellent teachers in the United States. *Bulletin of the Council for Research in Music Education*, *132*, 51-84.

Gardner, Howard. (1993). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.

Goh, W., & Pisoni, D. (2003). Effects of lexical competition on immediate memory span for spoken words. *The Quarterly Journal of Experimental Psychology: Section A*, *56*, 929-954.

Green, S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534-537.

Harris, P., & Silberstein, R (1999). Steady-state visually evoked potential (SSVEP) responses correlate with musically trained subjects' encoding and retention Phases of musical working memory task performance. *Australian Journal of Psychology*, *51*, 140-146.

Ho, Y.C., Cheung, M.C. & Chan, A.S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*, 439-450.

House, A., Williams, C., Hecker, M., and Kryter, K. (1965). Articulation-testing methods: consonantal differentiation with a closed-response set. *Journal of the Acoustical Society of America* 37, 158-166.

Huntsinger, C., & Jose, P. (1991). A test of Gardner's modularity theory: A comparison of short-term memory for digits and tones. *Psychomusicology*, *10*, 3-18.

Jakobson, L., Cuddy, L. & Kilgour, A. (2003). Time tagging: A key to musicians' superior memory. *Music Perception*, *20*, 307-313.

Kail, R., & Hall, L.K. (2001). Distinguishing short-term memory from working memory. *Memory and Cognition*, *29*, 1-9.

Karpicke, J. & Pisoni, D.B. (in press). Using immediate memory span to measure implicit learning. *Memory and Cognition.*

Kilgour, A., Jakobson, L., & Cuddy, L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory and Cognition*, *28*, 700-710.

Koelsch, S., Gunter, T., Cramon, D., Zysset, S., Lohmann, G., & Friederici, A. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage*, *17*, 956-966.

Lamb, S., & Gregory, A. (1993). The relationship between music and reading in beginning readers. *Educational Psychology*, *13*, 19-28.

LaPointe, L.B., & Engle, R.W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 1118-1133.

Lewellen, M., Goldinger, S., Pisoni, D., & Greene, B. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, *3*, 316-330.

Luce, P., and Pisoni, D. (1998). Recognizing spoken words: the neighborhood activation model. *Ear and Hearing* 19, 1-36.

Maess, B., Koelsch, S., Gunter, T., & Friederici, A. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, *4*, 540-545.

Marin, O., & Perry, D. (1999). Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press.

McMahon, O. (1982). A comparison of language development and verbalization in response to auditory stimuli in pre-school age children. *Psychology of Music*, *12*, 94-105.

Messerli, P., Pegna, A., & Sordet, N. (1995). Hemispheric dominance for melody recognition in musicians and non-musicians. *Neuropsychologia*, *9*, 97-113.

Munzer, S., Berti, S., & Pechmann, T. (2002). Encoding of timbre, speech, and tones: Musicians vs. non-musicians. *Psychologische Beitrage*, *44*, 187-202.

Nusbaum, H., Pisoni, D., & Davis, C. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10* (pp. 357-376). Bloomington, IN: Speech Research Laboratory, Indiana University.

Owen, A. (2000). The role of the lateral frontal cortex in mnemonic processing: The contribution of functional neuroimaging. *Experimental Brain Research*, *133*, 33-43.

Owen, A., Lee, A., & Williams, E. (2000). Dissociating aspects of verbal working memory within the human frontal lobe: Further evidence for a "process-specific" model of lateral frontal organization. *Psychobiology*, *28*, 146-155.

Patel, A., Gibson, E., Ratner, J., Besson., M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neurosci*ence, *10*, 717-33.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory and Cognition*, *20*, 314-320.

Penhue, V.B., Zatorre, R., & Evans, A. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, *19*, 752-765.

Pisoni, D.B. & Cleary, M. (2004). Learning, memory and cognitive processes in deaf children following cochlear implantation. In F.G. Zeng, A.N. Popper & R.R. Fay (Eds.), *Springer Handbook of Auditory Research: Auditory Prosthesis, SHAR Volume X*. Pp. 377-426.

Rauscher, F.H., Shaw, G.L, & Ky, K.N. (1993). Music and spatial task performance. *Nature, 365*, 611.

Saito, S., & Ishio, A. (1998). Rhythmic information in working memory: Effects of concurrent articulation on reproduction of rhythms. *Japanese Psychological Research*, *40*, 10-19.

Salamé, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology, 41*, 107-122.

Schellenberg, E.G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*, 511-514.

Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267(5198)*, 699-701.

Schlaug, G., Jancke, L., Huang, Y., Staiger, J., & Steinmetz, H. (1995b). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*, 1047-1055.

Schneider, P., Scherg, M., Dosch, H., Specht, H., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*, 688-694.

Sommers, M., Kirk, K., & Pisoni, D. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing*, *18*, 89-99.

Stallings, L., Kirk, K., Chin, S, & Gao, S. (2002). Parent word familiarity and the language development of pediatric cochlear implant users. *The Volta Review*, *102*, 237-258.

Stewart, L., Walsh, V. & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66*, 183-195.

Tervaniemi, M., Medvedev, S., Alho, K., Pakhomov, S., Roudas, M., Van Zuijen, T., & Naatanen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*, *10*, 74-79.

Texas Instruments. (1991). TI 46-word speaker-dependent isolated word corpus (CD-ROM). Gaithersburg: NIST.

Thomas, E., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305-307.

Thompson, W., Schellenberg, G. & Husain, G. (2001). Arousal, mood, and the Mozart Effect. *Psychological Science*, *12*, 248-251.

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). *The Stanford-Binet Scale of Intelligence*. Riverside: Chicago.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children--Third Edition*. San Antonio, TX: Psychological Corporation.

Zatorre, R., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, *114*, 2403-2417.