The Working Memory of Musicians and Nonmusicians

Francesca Talamini, Barbara Carretti, &
Massimo Grassi
University of Padova, Padova, Italy

Musicians have superior performances compared to nonmusicians in many auditory perception tasks. This superiority extends to memory tasks such as the digit span. Literature suggests that the musicians' advantage unfolds along two axes: sensory modality (musicians perform better when the task is auditory) and task complexity (musicians tend to perform better in the forward and not — for example — backward digit span). In addition, it is unclear whether there are specific music abilities linked with improved performance in the digit span. Here, musicians and nonmusicians performed a digit span task that was presented aurally, visually, or audiovisually. The task was performed with or without a concurrent task (i.e., articulatory suppression) in order to explore the role of rehearsal strategies and also manipulate task complexity. Finally, music abilities of all participants were assessed using the Profile of Music Perception Skills (PROMS) test. Musicians had larger spans than nonmusicians regardless of the sensory modality and the concurrent task. In addition, the auditory and audiovisual spans (but not visual) were correlated with one subscale of the PROMS test. Findings suggest a general advantage of musicians over nonmusicians in verbal working memory tasks, with a possible role of sensory modality and task complexity.

Received: July 25, 2015, accepted March 8, 2016.

Key words: verbal working memory, musicians, nonmusicians, digit span, articulatory suppression

The skills of musicians are often surrounded by an aura of mysticism. However, extensive music training (i.e., such as that required to become a musician) does shape behavior in several ways. The most evident empirical result is that musicians outperform nonmusicians in a variety of music perception tasks. Musicians are, for example, more proficient than nonmusicians in recognizing melodies presented in transposition (Halpern, Bartlett, & Dowling, 1995) or at a faster or slower tempo (Andrews, Dowling, Bartlett, & Halpern, 1998), and they are better in detecting mistuned notes (Koelsch, Schroger, & Tervaniemi, 1999; Schellenberg & Moreno, 2010). In addition, the effect of music training on auditory perception is not limited to music stimuli. Musicians, for example, outperform nonmusicians in classic psychoacoustical tasks such as frequency (Micheyl, Delhommeau, Perrot, & Oxenham, 2006) or temporal discrimination (Rammsayer & Altenmüller, 2006). Noticeably, the auditory ability of musicians also extends to speech: they are, in fact, better at recognizing the prosody of a sentence (Deguchi et al., 2012; Schön, Magne, & Besson, 2004) or understanding speech in noise (Parbery-Clark, Skoe, & Kraus, 2009).

The superior skills of musicians are not restricted to the auditory abilities. In fact, music training has an impact on several cognitive processes, such as working memory (WM). According to Baddeley and Hitch's (1974) classic model, WM consists of a central executive system that controls attentional resources and two subsidiary systems, the phonological loop and visuospatial sketchpad. While the phonological loop stores and manipulates verbal information (but also tonal stimuli; see Salamé & Baddeley, 1989), the visuospatial sketchpad stores and manipulates visual and spatial information.

In relation to Baddeley's model, the superiority of musicians over nonmusicians can be observed in tasks involving the phonological loop; for example, musicians outperform nonmusicians in WM tasks that use music stimuli (e.g., Pallesen et al., 2010; Schulze, Dowling, & Tillman, 2012; Schulze, Mueller, & Koelsch, 2011; Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011) and verbal material such as digits (George & Coch, 2011; Hansen, Wallentin, & Vuust, 2012; Lee, Lu, & Ko, 2007; Ramachandra, Meighan, & Gradzki, 2012). In contrast, results in visuospatial tasks (i.e., tasks that tap the visuospatial sketchpad) are mixed (see Amer, Kalender, Hasher, Trehub & Wong, 2013; Hansen et al., 2012; Lee et al., 2007). Furthermore, musicians tend to perform better in simple verbal span tasks that require maintenance (e.g., forward digit span, see Hansen et al., 2012) and, to a lesser extent, in complex verbal span tasks that require maintenance and manipulation (such as the backward digit span or the operation span task) or maintenance and the execution of a concurrent task, e.g., digit span with articulatory suppression (see

Music Perception, volume 34, issue 2, pp. 183–191, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2016.34.2.183

Franklin et al., 2008; George & Coch, 2011; Hansen et al., 2012; Nutley, Darki, & Klinberg, 2014).

For example, Hansen et al. (2012) investigated the WM abilities of expert musicians, amateur musicians, and nonmusicians by presenting the forward digit span (presented auditorily) of the Wechsler Adult Intelligence Scale (WAIS III; Wechsler, 1997a), spatial span of the Wechsler Memory Scale (WMS-III; Wechsler, 1997b), and Musical Ear Test (MET), which measures music abilities (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). Music abilities refer to those perceptual skills that do not pertain exclusively to musicians but can be found in individuals who have not received music training. Examples of these abilities include the ability to discriminate pitch, timbre, melodies, rhythms, and so on (i.e., those "natural music abilities or the innate potential to succeed as a musician"; Schellenberg & Weiss, 2013, p. 499). Hansen et al. (2012) observed that expert musicians demonstrated better WM performance than nonmusicians in the forward but not backward digit span, suggesting that music training improves WM in simple but not complex tasks. In contrast, no difference between groups was observed both in the forward and backward spatial spans. Finally, performance in the MET (in particular, in the rhythm subtest) was correlated with forward digit span scores.

Franklin et al. (2008) investigated WM and long-term memory of musicians and nonmusicians. They used the reading and operation spans, and observed a superiority of musicians over nonmusicians in both tasks. Long-term memory was investigated using the Rey Auditory Verbal Learning Test (RAVLT). To limit rehearsal strategies, the RAVLT was performed with and without articulatory suppression. Musicians' performance was better only when the task was performed without articulatory suppression. The authors hypothesized that the superiority of musicians in memory tasks could be linked to enhanced rehearsal strategies.

In contrast, a larger effect of music training was reported by Lee et al. (2007) in children. The authors investigated the effect of music training in verbal and visuospatial WM tasks in children and adults. The authors tapped the verbal domain via the forward and backward digit spans, nonword span, and operation span, and the visuospatial domain via the spatial span. Lee et al. found that the effect of music training on different components of WM depends on age. Children with music training performed better than children without training in both complex and simple tasks (but see Ho, Cheung, & Chan, 2003). In contrast, adults with music training performed better than adults without training only in the forward digit span and in the nonword span tasks. The latter result supports a relationship between music training and improved verbal WM in adults. However, this improvement does not seem to extend to more complex tasks such as the backward digit span and operation span.

Tierney, Bergeson-Dana, and Pisoni (2008) investigated the role of presentation modality on the superior performance of musicians over nonmusicians in verbal WM tasks. The authors investigated whether musicians showed an advantage when stimuli were presented auditorily. In this study, musicians and nonmusicians (gymnasts, psychology undergraduate students, and video games players) were compared on a modified version of the Simon Memory Game. In the visual modality, the four colored panels of a box were illuminated one at time in random order and each participant was asked to reproduce the sequence of colors by pressing the appropriate panels on the box. In the auditory modality, each participant listened to the names of the colors one at time randomly and was asked to reproduce the sequence by pressing the panels of the box. In the audiovisual modality the panels were illuminated while each participant listened to the name of the corresponding color. Musicians performed better than nonmusicians only in the auditory condition. No differences between groups emerged in the other conditions. This result suggests that musicians may perform better on WM tasks when stimuli are delivered auditorily.

Overall, literature suggests that musicians are better than nonmusicians in verbal WM tasks delivered auditorily (e.g., Hansen et al., 2012; Ho et al., 2003; Tierney et al., 2008). However, only Tierney et al. manipulated modality using one task-type (the Simon memory game). Therefore, it is possible that the superiority of musicians over nonmusicians is due to a better processing of auditory stimuli rather than to a better processing of verbal material per se (but see also Schulze et al., 2011, and Williamson, Baddeley, & Hitch, 2010, who did not observe any difference). In addition, literature reveals that musicians' performance is better in simple rather than complex tasks (e.g., Lee et al., 2007). In particular, it is possible that the superior performance of musicians in simple tasks is due to enhanced rehearsal strategies, as suggested by Franklin et al. (2008). Note that, in contrast, results regarding music perception tasks (such as pitch discrimination) often show the opposite pattern, i.e., greater differences between musicians and nonmusicians when the task is difficult rather than easy (e.g., Schön et al., 2004). Furthermore, literature reveals a relatively unexplored issue: only some music abilities could be correlated with good WM performance (Hansen et al., 2012; Wallentin et al., 2010) and these abilities in particular might explain the better WM performance of musicians compared to nonmusicians. However, to the best of our knowledge, only two studies (Hansen et al.; Wallentin et al.) assessed the music abilities of musicians and nonmusicians using an independent test (MET). In all other studies, the participants were simply divided into two groups: musicians and nonmusicians. Treating participants (i.e., musicians and nonmusicians) as a continuum for individual music abilities might clarify which ability explains differences in verbal WM performances.

The aim of the current study was to investigate the role of modality (auditory, audiovisual, and visual) and task complexity (in terms of executive processing demands) in the verbal WM of musicians and nonmusicians. We aimed to study whether the superiority of musicians in verbal WM tasks was limited to stimuli presented auditorily or extended to other presentation modalities (Tierney et al., 2008). As far as we know, all the studies that have compared the digit span scores of musicians vs. nonmusicians used auditory stimuli (i.e., the researcher reads the digits to each participant). Therefore, we wanted to investigate whether musicians performed better than nonmusicians when digits were presented visually (e.g., each participant reads the digits on a computer screen) and audiovisually.[1] The classic digit span task responded to these needs: digits can be presented auditorily, visually, and audiovisually. In order to investigate whether the difference in the digit span between groups is dependent on the phonological loop (perhaps due to more efficient rehearsal strategies; Franklin et al., 2008) and to manipulate task complexity (e.g., Hansen et al., 2012), participants were asked to perform the digit span with a concurrent task (articulatory suppression), which blocks the phonological loop by inhibiting verbal rehearsal (Baddeley, Thomson, & Buchanan, 1975) and enhances complexity in terms of cognitive load. We also investigated whether WM performance differed depending on music ability, as assessed using the test Profile of Music Perception Skills (PROMS, Law, & Zentner, 2012).

Method

PARTICIPANTS

Thirty-six students, all native Italians, participated voluntarily in the study. Participants provided written, informed consent to participate in the experiment. Eighteen participants (eight females, mean age = 22.0 years, mean education = 14.9 years) were trained musicians, defined as being a music conservatory student with a minimum of 7 years of training. Musicians received a mean of 12.3 years of music training (range = 7-26) and practiced their instrument an average of 19.6 hours per week. The nonmusicians group (fourteen females, mean age = 23.2 years, mean education = 16.7 years) was defined as having received no music training beyond the basic music classes of the Italian Middle School curriculum. In addition, because musicians were unfamiliar with psychological experiments (none had participated in a psychology experiment before) the nonmusicians were also selected with the same criterion. All participants had audiometric thresholds below 20 dB HL for frequencies 500, 1500, and 4000 Hz, and normal (or corrected to normal) vision.

Participants were asked to complete two WAIS-IV (Wechsler, 2008) subtests, the Vocabulary and the Visual Puzzles, to control for general verbal and nonverbal cognitive abilities. A set of independent samples *t*-tests were used to evaluate whether musicians and nonmusicians differed in age, educational level, and Vocabulary and Visual Puzzle scores. Nonmusicians had a higher educational level than musicians, $t(34) = 3.17$, $p = .003$, but the two groups did not differ with regard to age, $t(34) = 1.21$, $p = .24$, Visual Puzzles score, $t(34) = 0.50$, $p = .61$, and Vocabulary score, $t(34) = 1.11$, $p = .27$ (see Table 1).

APPARATUS

The digit span test and the PROMS were administered using an ASUS computer (Cpu Intel i5 650 3.20 GHz, Motherboard Asus P7H55-V RAM 4 GB, Graphic Card AMD Radeon HD 5700 Series, OS Windows 7 Professional 64 bit). The computer was connected to a monitor (NEC MultiSync FE950+) and M-AUDIO FastTrack Pro sound card. The output of the sound card was delivered to a pair of Sennheiser HD 580 headphones. Each participant took the computer tests seated inside

TABLE 1. *Age, Education, Performance (Raw Scores) in the WAIS-IV Visual Puzzle and Vocabulary Subtests*

	Musicians		Nonmusicians	
	M	*SD*	*M*	*SD*
Age (years)	22.06	3.80	23.28	2.02
Education (years)	14.94	2.15	16.78	1.17
Visual Puzzle (max score 26)	17.39	5.37	18.17	3.75
Vocabulary (max score 57)	44.72	7.75	42.17	5.82

[1] The audiovisual condition was originally included to test whether in a WM task the integration of unisensory stimuli followed the same rules observed in perceptual tasks (Ernst & Banks, 2002; Grassi & Pavan, 2012); namely, better performance and smaller variance in the audiovisual condition compared to the auditory only and visual only conditions.

a single walled IAC sound proof booth. The digit span test was administered using a custom-coded MATLAB program with the Psychophysics Toolbox extensions (Kleiner et al., 2007). The PROMS test was administered through its website. The auditory stimuli for the auditory digit span test were recorded by a male speaker (the last author) with neutral prosody using a Shure SM 58 microphone. Single digit recordings were edited using the CoolEdit Pro software (Syntrillium Software) and assembled in sequences by the custom made MATLAB program that also generated the visual and the audiovisual stimuli.

MATERIAL

Digit span test. During the digit span test, participants saw, heard, or saw and heard (depending on the modality) a random sequence of numbers (1 to 9) and were instructed to input the sequence in the correct order via the computer's numeric keypad. Each sequence was preceded and followed by a 500 Hz, 500 ms pure tone (auditory modality) or 500 ms long asterisk (visual modality) that signalled the beginning and end of the trial. Both the tone and asterisk were presented in the audiovisual modality. Across modalities, the numbers were presented at a pace of one every 1.5 seconds. In the auditory version of the span-test, the pace was calculated as the temporal distance between the tonic accents of two consecutive numbers (e.g., the distance between /u/ and /e/ in the "uno – tre" sequence, respectively one and three in English). In the visual version, the pace was calculated as the temporal distance between the onsets of two consecutive numbers. In the visual version of the span-test, the digits were written in Arabic numbers (24 points Arial font) centered on the computer screen for 750 ms. In the audiovisual span-test, the auditory and visual digits were presented simultaneously. The test began with a sequence of three numbers. There were two sequences for each level of difficulty. After one correct response (i.e., all numbers of one or both sequences given in the correct order) the sequence length was increased by one number. This procedure was iterated until the participant gave wrong answers to both sequences. Successively, participants performed the test with articulatory suppression. In this condition, participants were instructed to perform the digit span test while repeating loudly and quickly "la la la" during the presentation of the sequence. Participants were instructed to begin and end the articulatory suppression simultaneously with the tones (or asterisks) signalling the beginning and end of the sequence. Outside the testing booth, the experimenter assessed that the participant performed the concurrent task. In all modalities (i.e., including the visual modality) participants wore headphones and were asked to look at the center of the screen while the sequence was presented. Therefore, participants could not look at the keys of the numeric keypad to spatially encode the sequences. The experimenter assessed this from the outside of the booth through the booth's window.

Musical abilities - mini PROMS (Law & Zentner, 2012). This test consisted of four sections (melody, tuning, speed, and beat) performed in sequence by the participant. Each section was articulated in ten trials. The structure of the trials was identical in the four sections. Each participant listened to a standard stimulus twice. The standard stimulus was followed by a comparison stimulus that could be identical or different from the standard. Participant reported whether the comparison was different (or not) from the standard on a five points scale answer (i.e., *definitely same*, *probably same*, *I don't know*, *probably different*, *definitely different*). The correct answers chosen with maximum confidence (definitely same or different) scored two points. The less confident correct answers (probably same or different) scored one point. The wrong and *I don't know* answers scored zero points.

In the melody section, the standard and comparison stimuli were melodies of constant rhythm. The comparison could be either identical to the standards or different in one (or more) notes. The difficulty of the trials was manipulated by increasing note density (i.e., number of notes per time unit) and atonality. In the tuning, the standard and comparison stimuli were chords (the notes C4, E4, G4, and C5) of 1.5 s duration. The comparison could be either identical to the standards or different in the note E that could be shifted from 10 to 50 cents. In the speed section, the standard and comparison stimuli could be either a synthetic rhythmic structure or a recorded music sample. The comparison could be either identical to the standards (i.e., identical beats per minute, BPM) or different in speed that could be varied between +1 BPM to +7 BPM. In the beat section, standards and comparison were rhythmic patterns of clicks. The rhythm was created by giving an accent (i.e., a 3 dB increment) to a subset of the clicks. In the easy trials, intensity changes were applied to most sound events so as to increase the probability of detecting the alteration. In the moderate and difficult test trials, there were fewer intensity changes, which required more subtle perceptual skills to be identified.

PROCEDURE

Participants signed the informed consent form engaged in the audiometric screening prior to the experiment. The experimenter then collected demographic details

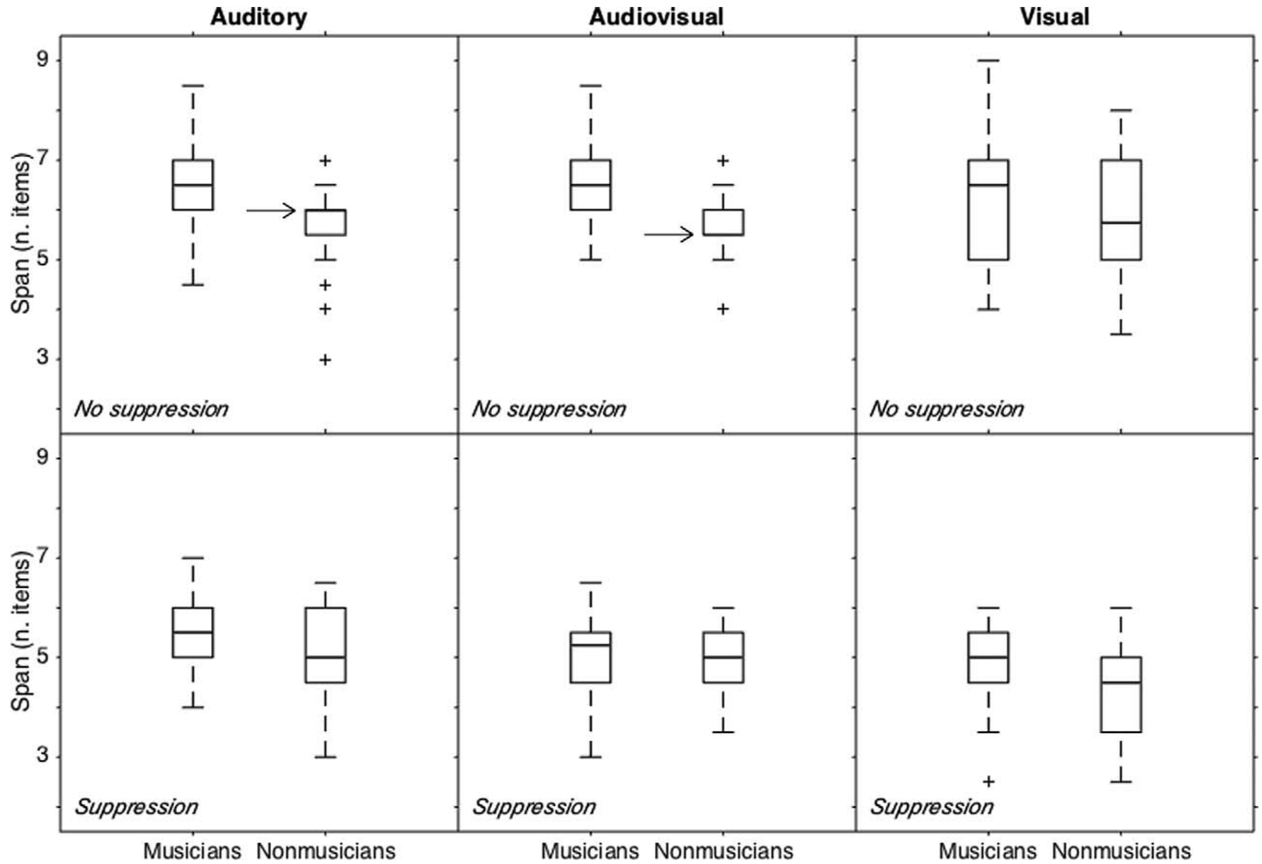

FIGURE 1. Digit-span of the participants in the various modalities and conditions of the study. In each box, the median is indicated by horizontal lines inside the box (or right-facing arrows in the two incidences where the median is the same as the top or bottom boundary of a box). The edges of the box are the 25th and 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50% and symbols are outliers.

and participants performed the Vocabulary task and digit span tasks. Participants took the different conditions of the digit span task in different order. Half of participants of each group started with the no-articulatory suppression condition whereas the other half started with the articulatory suppression. The modality order (i.e., auditory, visual, and audiovisual) was counterbalanced within each group. After the digit span task, participants completed the Visual Puzzles task and the PROMS test. Finally, participants completed a questionnaire on their music habits. Musicians answered to a few further questions about their music training. The experimental session lasted about one hour.

Results

Span measure of each participant in the various modalities and conditions (see Figure 1) was computed and used as the dependent variable in a two-ways ANOVA with two within-subjects factors (Modality = auditory vs. visual vs. audiovisual, Suppression = with articulatory suppression vs. without articulatory suppression) and one between-subjects factor (musicians vs. nonmusicians). The ANOVA revealed that musicians overall performed better than nonmusicians, $F(1, 34) = 4.41$, $p = .04$, $\eta_p^2 = .12$. The performance was worse in the suppression compared to no suppression condition, $F(1, 34) = 94.22$, $p < .001$, $\eta_p^2 = .74$; however, this result did not interact with group, $F(1, 34) = 1.98$, $p = .17$, $\eta_p^2 = .06$. The modality factor approached significance, $F(2, 68) = 2.81$, $p = .07$, $\eta_p^2 = .08$, but did not interact with group, $F < 1$. The interaction between suppression and modality was significant, $F(2, 68) = 4.09$, $p = .02$, $\eta_p^2 = .11$, but did not interact with group, $F(2, 68) = 1.47$, $p = .24$, $\eta_p^2 = .04$.

Although we did not observe an interaction between sensory modality and group, we explored the modality hypothesis (Tierney et al., 2008) and computed three separate ANOVAs with one within-participants factor (with vs. without articulatory suppression) and one

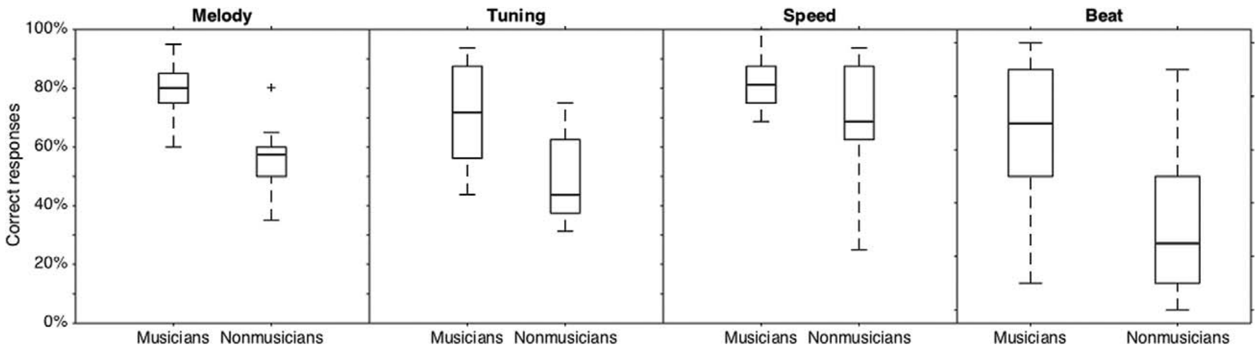

FIGURE 2. Boxplots representing the performance of the participants in the various subscales of the PROMS test. In each box, the median is indicated by horizontal lines inside the box. The edges of the box are the 25th and 75th percentiles. The whiskers are the interquartile range (i.e., Q3-Q1) augmented by 50% and symbols are outliers.

between-participants factor (i.e., group) separately for auditory, visual, and audiovisual spans. Auditory span was larger for musicians than for nonmusicians, $F(1, 34) = 5.06$, $p = .03$, $\eta_p^2 = .13$. The dimension of the difference was large ($d = .80$) according to Cohen's (1988) guidelines. In addition, both groups showed a larger span without articulatory suppression than with articulatory suppression: $F(1, 34) = 28.68$, $p < .001$, $\eta_p^2 = .46$. The interaction between group and condition was not significant, $F(1, 34) = 1.14$, $p = .29$, $\eta_p^2 = .03$. Audiovisual span was larger for musicians than nonmusicians, $F(1, 34) = 4.19$, $p = .02$, $\eta_p^2 = .11$, and larger without articulatory suppression than with it, $F(1, 34) = 48.06$, $p < .001$, $\eta_p^2 = .59$. In addition, the interaction between group and condition was significant, $F(1, 34) = 6.58$, $p = .02$, $\eta_p^2 = .17$. Post hoc analyses revealed that musicians had a larger span than nonmusicians in the audiovisual modality with no articulatory suppression ($p = .005$), with a large effect size ($d = 1.01$), but not with it ($p = .63$). Both groups, however, showed better performance in the no suppression condition with respect to that with articulatory suppression (musicians, $p < .001$, nonmusicians, $p = .004$). Finally, visual spans were not different for the two groups, $F(1, 34) = 1.35$, $p = .25$, $\eta_p^2 = .04$, and they were larger without articulatory suppression, $F(1, 34) = 36.92$, $p < .001$, $\eta_p^2 = .52$. The interaction between group and condition was not significant, $F < 1$.

The PROMS scores (see Figure 2) were subjected to four Bonferroni-adjusted independent samples t-tests. The t-tests revealed that musicians outperformed nonmusicians in all the PROMS subscales, i.e., melody, $t(34) = 6.09$, $p < .0001$ ($d = 2.03$), tuning, $t(34) = 4.20$, $p = .0002$ ($d = 1.40$), speed, $t(34) = 2.50$, $p = .02$ ($d = .83$), and beat, $t(34) = 3.90$, $p = .0004$ ($d = 1.31$).

TABLE 2. *Correlations Between the WM Span For Various Modalities and Conditions and the PROMS Subscales*

	Melody	Tuning	Speed	Beat
Span Auditory	.48*	.22	.28	.06
Span Audiovisual	.51*	.35	.39	.22
Span Visual	.16	.34	.17	.01
Span Auditory (S)	.36	.23	.26	.11
Span Audiovisual (S)	.10	.24	.31	.01
Span Visual (S)	.37	.30	.27	.15

Note: Rows labels followed by (S) highlight the span with suppression. *$p < .05$

Finally, 24 FDR-adjusted (Benjamini & Hochberg, 1995) correlations were run in order to evaluate whether any of the PROMS subtests related to span scores. The correlations are reported in Table 2.

Significant correlations were observed between the melody subtest and the span task in the auditory modality, $r(36) = .48$, $p = .04$, and the audiovisual modality, $r(36) = .51$, $p = .02$, with no articulatory suppression. In other words, the better the participant performed in the melody subtest, the better s/he performed in the digit span when the modality was auditory and audiovisual only when there was no articulatory suppression. In contrast, the melody score did not correlate with the span score in the visual modality, $r(36) = .16$, $p > .05$. Figure 3 shows the correlation graphs highlighting the relationship between the score at the melody subtest and the digit span in the auditory and audiovisual modalities. Finally, we also calculated twelve FDR-adjusted correlations between years of music training and span measures and between hours of weekly practice and span measures for the musicians group only. We observed a moderate positive correlation between years of music training and spans in the audiovisual modality with no articulatory suppression, $r(18) = .48$, and

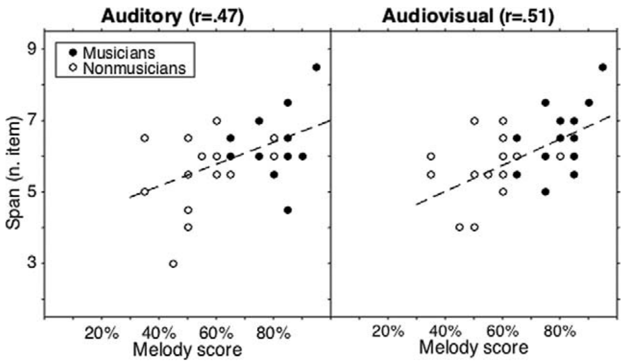

FIGURE 3. Correlational plots. Auditory and audiovisual digit span as a function of the melody score in the PROMS test. Each symbol represents one participant. The graph title includes the value of the correlation coefficient.

a moderate negative correlation between hours of weekly practice and spans in the auditory modality with articulatory suppression, $r(18) = -.43$. However, both correlation were not significant when the p value was FDR-adjusted.

Discussion

The current study investigated the performance of musicians and nonmusicians in a verbal WM task. The investigation developed along two main axes: 1) whether musicians showed better performances than nonmusicians when the task was delivered auditorily compared to other modalities (e.g., Hansen et al., 2012; Tierney et al., 2008); and 2) whether musicians showed better performances than nonmusicians when the task was complex and they could not rely on verbal rehearsal strategies (e.g., Franklin et al., 2008). In addition, we investigated which individual music ability might be linked to a better verbal WM performance. In order to investigate these issues, musicians and nonmusicians were asked to perform a digit span task delivered auditorily, visually, or audiovisually. In order to investigate the role of rehearsal strategies and task complexity, participants were asked to perform the digit span task with or without articulatory suppression. Finally, several music abilities were assessed via an independent test (i.e., PROMS) and were correlated with the span.

Results revealed that musicians performed better than nonmusicians in the digit span task, regardless of the presence of a concurrent task. Articulatory suppression did not affect group differences. This result suggests that the advantage of music training can be observed in maintenance tasks (i.e., simple digit span) as well as in complex tasks (i.e., digit span with a concurrent task). Previous studies have shown mixed results, with some showing a general strength in the verbal component of WM, both involving passive storage and active processing (e.g., Franklin et al., 2008), and others showing specific strength only in maintenance components (e.g., Hansen et al., 2012; Lee et al., 2007). Our results seem to support the former rather than the latter findings.

Presentation modality did not affect the differences between groups. In fact, results revealed that musicians performed better regardless of the modality. Literature suggests that musicians perform better when stimuli are presented auditorily. However, only Tierney et al. (2008) and the current study investigated modality with a single task-type. Here, although the first ANOVA did not reveal an interaction between modality and group, the second set of ANOVAs (run to test directly the modality hypothesis) suggested that the advantage of the musicians might be more prominent in the auditory and audiovisual modalities. In the auditory modality, indeed, the superior performance was observed with and without articulatory suppression. Interestingly, in the audiovisual modality, musicians showed better performances than nonmusicians without articulatory suppression. However, differences between the groups disappeared when the digit span task was presented with a concurrent task. The superior ability of the musicians in the audiovisual condition might be explained in two ways. On the one hand, the superior ability of musicians may be linked to the fact that music training helps integrate information coming from different sensory modalities (e.g., see Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012; Paraskevopoulos, Kraneburg, Herholz, Bamidis, & Pantev, 2015). On the other hand, it is possible that musicians relied mostly on the auditory modality. In synthesis, the modality hypothesis was weakly supported.

Another aim of the current study was to understand whether better WM performance was associated with specific music abilities. Here, auditory and audiovisual performances without articulatory suppression were found to be correlated with melody subtest. This correlation can perhaps be explained by the nature of the melody task. This subtest required participants to remember a melody and compare it to a second one. The difficulty of the task was manipulated by increasing the number of notes of the melody (i.e., as in the span task). It is evident that memory has a great role in this task. Interestingly, however, there was no correlation between the visual digit span and the "melody" subtest and this implies that the modality through which stimuli are presented (i.e., auditory vs visual) should not be ignored (Tierney et al., 2008). Noticeably, Hansen et al.

(2012) observed a correlation between the forward digit span and the rhythm MET subtest (Wallentin et al., 2010). In this subtest the difficulty is modulated by increasing the number of tones the participant has to listen to (i.e., 4 to 11). Hansen et al. (2012), however, did not observe a correlation between the melody subtest of MET and the span performance.

Finally, one can argue whether the superior WM performance of musicians is due to the fact that they have better general cognitive abilities than nonmusicians (e.g., Schellenberg, 2004). Here, the results of the two WAIS-IV subtests (i.e., Visual Puzzles, Vocabulary) did not support this hypothesis: we did not observe any significant difference between musicians and nonmusicians in these two tasks.

In summary, current results support the superiority of musicians in verbal WM, regardless of the complexity of the task and modality. However, modality might play a role as suggested by literature and by our second set of ANOVAs. Some music abilities were correlated with auditory and audiovisual span (but not with the visual one), suggesting that specific music skills might be in a relationship with the digit span.

Author Note

The present work was performed with no funding. The authors thank Marcel Zentner for providing the PROMS test. We also thank reviewers for the fruitful comments.

Correspondence concerning this article should be addressed to Massimo Grassi. Department of General Psychology, University of Padova, via Venezia 8, 35131, Padova, Italy. E-mail: massimo.grassi@unipd.it

References

Amer, T., Kalender, B., Hasher, L., Trehub, S. E., & Wong, Y. (2013). Do older professional musicians have cognitive advantages. *PloS One*, *8*(8), e71630. DOI: 10.1371/journal.pone.0071630

Andrews, M. W., Dowling, W. J., Bartlett, J. C., & Halpern, A. R. (1998). Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychology and Aging, 13*, 462-471. DOI: 10.1037/0882-7974.13.3.462

Baddeley, A. D., & Hitch. G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 8., pp. 47-89). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behaviour, 14,* 575-589. DOI: 10.1016/S0022-5371, 80045-4

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistic Society B*, *57*, 289-300.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schön, D., & Colombo, M. (2012). Sentence pitch change detection in the native and unfamiliar language in musicians and nonmusicians: Behavioral, electrophysiological and psychoacoustic study. *Brain Research, 1455*, 75-89. DOI: 10.1016/j.brainres.2012.03.034

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429-433. DOI:10.1038/415429a

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429-433. DOI: 10.1038/415429a

Franklin, M. S., Moore, K. S., Yip, C. Y., Jonides, J., Rattray, K., & Moher. J. (2008). The effects of musical training on verbal memory. *Psychology of Music, 36*, 353-365. DOI: 10.1177/0305735607086044

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083-1094. DOI: 10.1016/j.neuropsychologia.2011.02.001

Grassi, M., & Pavan, A. (2012). The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception, and Psychophysics, 74*, 1321-1333. DOI: 10.3758/s13414-012-0324-x

Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1995). Aging and expertise in the perception of musical transpositions. *Psychology and Aging, 10*, 325-342. DOI: 10.1037/0882-7974.10.3.325

Hansen, M., Wallentin, M., & Vuust, P. (2012). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, *41*, 779-793. DOI: 10.1177/0305735612452186

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. DOI: 10.1037/0894-4105.17.3.439

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*, 1-16.

Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport, 10*, 1309-1313.

Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PloS One, 7*, e52508. DOI: 10.1371/journal.pone.0052508

Lee, Y. S., Lu, M. J., & Ko, H. P. (2007). Effects of skill training on working memory capacity. *Learning and Instruction, 17*, 336-344. DOI: 10.1016/j.learninstruc.2007.02.010

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research, 219*, 36-47. DOI: 10.1016/j.heares.2006.05.004

Nutley, S. B., Darki, F., & Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Frontiers in Human Neuroscience*, *7*, 1-9 (article #926). DOI: 10.3389/fnhum.2013.00926

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PloS One*, *5*, e11120. DOI: 10.1371/journal.pone.0011120

Parbery-Clark, A., Skoe, E., & Kraus, E. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience, 29*, 14100-14107. DOI: 10.1523/JNEUROSCI.3256-09.2009

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Musical expertise induces audiovisual integration of abstract congruency rules. *The Journal of Neuroscience, 32*, 18196-18203.

Paraskevopoulos, E., Kraneburg, A., Herholz, S. C., Bamidis, P. D., & Pantev, C. (2015). Musical expertise is related to altered functional connectivity during audiovisual integration. *Proceedings of the National Academy of Sciences, 112*, 12522-12527.

Ramachandra, V., Meighan, C., & Gradzki, J. (2012). The impact of musical training on the phonological memory and the central executive: A brief report. *North American Journal of Psychology*, *14*, 541-548.

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception, 24*, 37-48. DOI: 10.1525/mp.2006.24.1.37

Salamé, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory, *Quarterly Journal of Experimental Psychology Section A, 41*, 107-122. DOI: 10.1080/14640748908402355

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*, 511-514. DOI: 10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G., & Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychology of Music, 38*, 209-221. DOI: 10.1177/0305735609339473

Schellenberg, E. G., & Weiss, M.W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 499-550). Amsterdam: Elsevier.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341-349. DOI: 10.1111/1469-8986.00172.x

Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. *Music Perception*, *29*, 255-267. DOI: 10.1525/mp.2012.29.3.255

Schulze, K., Müller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience*, *33*, 189-196. DOI: 10.1111/j.1460-9568.2010.07470.x

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping*, *32*, 771-783. DOI: 10.1002/hbm.21060

Tierney, A. T., Bergeson-Dana, T. R., & Pisoni, D. B. (2008). Effects of early musical experience on auditory sequence memory. *Empirical Musicology Review, 3*, 178-186.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test. A new reliable test for measuring musical competence. *Learning and Individual Differences, 20*, 188-196. DOI: 10.1016/j.lindif.2010.02.004

Wechsler, D. (1997a). *WAIS-III administration and scoring manual*. San Antonio. TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale* (3rd ed.). San Antonio. TX: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (WAIS–IV, 4th ed.). San Antonio. TX: NCS Pearson.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory and Cognition, 38*, 163-175.