Hearing Research 342 (2016) 112–123

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Musicians' edge: A comparison of auditory processing, cognitive abilities and statistical learning

Pragati Rao Mandikal Vasuki [a, b, c, *], Mridula Sharma [a, b], Katherine Demuth [a, b, c], Joanne Arciuli [b, d]

[a] Department of Linguistics, Australian Hearing Hub, 16 University Avenue, Macquarie University, New South Wales, 2109, Australia
[b] The HEARing CRC, 550 Swanston Street, Audiology, Hearing and Speech Sciences, The University of Melbourne, Victoria, 3010, Australia
[c] ARC Centre of Excellence in Cognition and its Disorders, Level 3, Australian Hearing Hub, 16 University Avenue, Macquarie University, New South Wales, 2109, Australia
[d] Faculty of Health Sciences, University of Sydney, 75 East St, Lidcombe, 1825, Australia

ARTICLE INFO

Article history:
Received 29 October 2015
Received in revised form
11 October 2016
Accepted 15 October 2016
Available online 19 October 2016

Keywords:
Musicians
Statistical learning
Auditory processing
Attention
Digit span

ABSTRACT

It has been hypothesized that musical expertise is associated with enhanced auditory processing and cognitive abilities. Recent research has examined the relationship between musicians' advantage and implicit statistical learning skills. In the present study, we assessed a variety of auditory processing skills, cognitive processing skills, and statistical learning (auditory and visual forms) in age-matched musicians ($N = 17$) and non-musicians ($N = 18$). Musicians had significantly better performance than non-musicians on frequency discrimination, and backward digit span. A key finding was that musicians had better auditory, but not visual, statistical learning than non-musicians. Performance on the statistical learning tasks was not correlated with performance on auditory and cognitive measures. Musicians' superior performance on auditory (but not visual) statistical learning suggests that musical expertise is associated with an enhanced ability to detect statistical regularities in auditory stimuli.

1. Introduction

Music is a quintessential multisensory activity and musical training involves engagement of multiple neural and cognitive resources. Musicians not only engage in auditory training due to many hours of listening and practising but also multimodal training involving reading and translation of complex symbolic notation into motor activity (Schlaug et al., 2005). Though it is difficult to differentiate abilities that prompt individuals to pursue music training from abilities that may result from music training, some cross-sectional studies comparing musicians with non-musician peers have shown that musicians perform better on certain auditory processing tasks (Fine and Moore, 1993; Micheyl et al., 2006; Zendel and Alain, 2012) and tasks of executive function (Bialystok and DePape, 2009; Zuk et al., 2014). In addition, neurophysiological and brain imaging studies have shown differences in the brain structure and function of musicians and non-musicians. For example, professional musicians have larger grey matter volume in primary motor and somatosensory areas, premotor areas, anterior superior parietal areas, and in the inferior temporal gyrus bilaterally (Gaser & Schlaug, 2003). The structural and functional changes associated with musical expertise are in line with an experience-dependent model of neuroplasticity (Münte et al., 2002). In summary, there is widespread interest in the musicians' advantage in various abilities. The current research was designed to explore whether auditory processing, cognitive processing, and implicit learning of statistical regularities – statistical learning – differ as a function of musical expertise.

1.1. Auditory processing and musical expertise

Some of the most widely investigated abilities associated with musical expertise pertain to auditory processing. Auditory processing is an umbrella term including spectral and temporal processing. Additionally, measures of temporal processing may include envelope processing and fine structure processing. The tests used to measure auditory processing include tests of frequency discrimination, discrimination of iterated rippled noise, detection of

* Corresponding author. Room 602, Level 1, Australian Hearing Hub, Macquarie University, Sydney, 2109, Australia.
E-mail address: pragati.mandikal-vasuki@mq.edu.au (P.R. Mandikal Vasuki).

http://dx.doi.org/10.1016/j.heares.2016.10.008
0378-5955/

amplitude modulation, detection of gaps in noise, dichotic listening tests, and perception of speech in noise.

When considering the relationship between musical expertise and auditory processing, two issues come into play. First, as far as we are aware, no previous study has compared the performance of musicians and non-musicians on auditory processing tasks that address both spectral and temporal processing abilities. Second, the extent to which musical expertise is associated with superior performance on auditory processing tasks remains a controversial issue. On one hand, research has consistently shown that musicians have better frequency discrimination than non-musicians (Kishon-Rabin et al., 2001; Micheyl et al., 2006). On the other hand, contradictory evidence exists for musicians' superior auditory skills for gap detection (Ishii et al., 2006; Rammsayer and Altenmüller, 2006; Zendel and Alain, 2012), perception of speech in noise (Parbery-Clark et al., 2009; Ruggles et al., 2014), dichotic listening tests (Nelson et al., 2003; Špajdel et al., 2007), and other temporal processing skills (Iliadou et al., 2014; Ishii et al., 2006). Given these inconsistencies in the literature, we incorporated a comprehensive battery of auditory processing tasks in the current study.

1.2. Cognitive abilities and musical expertise

Musical expertise might be associated with some aspects of cognitive processing. For instance, it has been reported that adults and children who have undertaken music training have better working memory as measured through digit span and non-word span (Lee et al., 2007). Better performance on executive function measures such as verbal fluency, design fluency and backwards digit span for musicians has also been reported (Zuk et al., 2014). Still, there have been some ambivalent results as to whether or not musical expertise is associated with enhanced cognitive abilities. Using a large battery of tasks assessing cognitive skills such as verbal comprehension, word fluency, mental rotation, closure, perceptual speed, reasoning, and verbal memory, Brandler and Rammsayer (2003) found significant group differences in only two tasks – verbal memory and reasoning. Similar results were reported in another study where musicians were found to have better performance in only two out of the thirteen primary cognitive abilities tested – flexibility of closure and perceptual speed (Helmbold et al., 2005). Additionally, it is unclear whether enhancements are seen only in the auditory modality, such as in auditory attention tasks (Strait and Kraus, 2011), or also in the visual modality, such as divided visual attention tasks (Rodrigues et al., 2007). Given these gaps in the literature, we incorporated a battery of cognitive processing tasks in the current study including a task that assessed both visual and auditory attention.

1.3. Statistical learning and musical expertise

A growing area of interest is musicians' ability to learn statistical regularities implicitly, known as statistical learning (SL). SL was described in a seminal study by Saffran et al. (1996). They showed that participants are able to extract statistical regularities from a continuous stream of individually presented stimuli using information about transitional probabilities. SL has been shown in auditory (aSL) and visual (vSL) modalities. It is thought that SL ability may contribute to key mental activities including musical appreciation, object recognition, and language acquisition (Arciuli and von Koss Torkildsen, 2012; Rohrmeier and Rebuschat, 2012).

Similar to language, music is highly structured and listeners are able to extract regularities from music (François and Schön, 2010). Whilst being unaware of the complex patterns of music, it is possible to implicitly acquire musical knowledge and use this implicit knowledge to form expectancies, and extract regularities from continuous events. Heightened sensitivity to these statistical regularities in continuous speech or non-speech streams may be partly explained by the shared and overlapping cortical regions for music and language (OPERA hypothesis; Patel, 2010, 2011). It could also be argued that musical competence, which is acquired through repeated practise and exposure, primes and sharpens musicians' intuition for performing implicit learning tasks (Rohrmeier and Rebuschat, 2012). In addition, musicians have may have enhanced processing of auditory stimuli (for a detailed review see François and Schön, 2014). For these reasons, it is interesting to study SL in musicians.

Using neurophysiological measures such as electroencephalography and magnetoencephalography, musicians have been shown to have enhancements in neurophysiological indices (such as N100 or N400) in auditory tasks involving the extraction of distributional cues (François et al., 2014; François and Schön, 2011; Paraskevopoulos et al., 2012; Schön and François, 2011). To date, only two studies have demonstrated an advantage for adult musicians in aSL using behavioural indices (Shook et al., 2013 using morse code; Skoe et al., 2013 using tone doublets). A report of improved SL in a longitudinal study of 8- year old children learning music as opposed to a control painting group suggests that there may be a causal link between musical training and SL (François et al., 2013). Although musical expertise has been associated with improved skills in the visual domain, such as enhanced recognition of visual patterns, also known as design learning (Jakobson et al., 2008), an investigation of musicians' vSL has not been undertaken previously.

Any demonstrable musicians' advantage in SL raises further questions as to whether such an advantage is accompanied by advantages in auditory processing or other cognitive skills. Though not directly investigated, enhanced statistical learning of morse code in musicians was attributed to enhanced temporal encoding and/or cognitive skills in musicians (Shook et al., 2013). However, as far as we are aware, this has not been investigated empirically. We used an array of auditory processing tasks and cognitive processing tasks as well as measures of both auditory and visual SL to explore these questions.

1.4. The current study

The primary aims of this research were to ascertain whether musicians and non-musicians perform differently on: a) tests of auditory processing, b) tests of cognition, and c) tests of SL (aSL and vSL). We hypothesized that musical expertise would be associated with better performance on at least some of the auditory processing and cognition measures. We also hypothesized that musicians might outperform non-musicians with regard to aSL but we were not sure what to expect with regard to vSL. Moreover, we were unsure whether performance on SL tasks would be related to performance on the auditory and cognitive tasks.

2. Methods

2.1. Participants

Musicians were defined as adults who started to learn/practise music before the age of 9 years and had at least 10 years of music playing/singing experience. This criterion is based on previous studies with similar populations (Ruggles et al., 2014; Strait et al., 2010). All musicians reported that they still actively practised music. Non-musicians had less than 3 years of musical experience. Eighteen musicians (5 males) and 22 non-musicians (5 males) participated in the study. There was no significant difference in the ages of the musicians (Mdn = 28.0) and non-musicians

(Mdn = 25.0) as assessed by a Mann-Whitney *U* test [U = 163, *p* = 0.35]. All participants were right handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and had normal or corrected to normal vision. At the time of data collection, four of the non-musicians were enrolled in the final year of a postgraduate program. Details about participants' music education, instruments played, and educational background are in Table 1.

The participants were recruited through advertisements distributed via group emails or announcement on Facebook community pages. All participants lived in the greater Sydney metropolitan area and had obtained an undergraduate degree. The study was approved by the Macquarie University Human Participants Ethics Committee and written consent was received from all participants. All participants were paid $40 for their participation.

2.2. Tests

All participants completed behavioural testing in a sound treated booth. Participants' hearing was screened at 15 dB HL (all octave frequencies from 0.5 to 8 kHz). Additionally, distortion product otoacoustic emissions (DPOAEs) and contralateral acoustic reflexes were present at clinically normal levels in all participants.

2.2.1. Auditory processing

The auditory processing tests were the dichotic digits test (3 pairs) (Strouse and Wilson, 1999), gaps in noise test (Baker et al., 2008), frequency discrimination of 1 kHz tones, threshold for discrimination of iterated rippled noise (Peter et al., 2014), threshold for detection of 4 Hz and 64 Hz amplitude modulation (Peter et al., 2014), and the listening in spatialized noise sentence test (LiSN-S, Cameron and Dillon, 2007). As test materials and scoring procedures have been published previously, only brief details of the administration and scoring procedures are mentioned in Table 2.

2.2.2. Cognition (memory, inhibition, and attention)

Working memory capacity was evaluated using forwards and backwards digit span subtests from the clinical evaluation of language fundamentals, 4th edition (CELF-IV; Semel et al., 2006). A Stroop colour word test was used to evaluate inhibition and selective attention. The integrated visual auditory continuous performance test was used to measure sustained attention in auditory and visual modalities (IVA-CPT, Turner and Sandford, 1995). A brief description of procedures for the tests is given in Table 2.

2.2.3. Statistical learning

SL was investigated unimodally in the auditory and visual domains. The separate aSL and vSL tasks were designed to be as similar as possible, using the embedded triplet paradigm with a familiarization phase followed by a separate test phase. The aSL and vSL tasks were adapted from Abla et al. (2008), Abla and Okanoya (2009), and Arciuli and Simpson (2011).

Table 1
Details of all participants' musical and educational background.

No	Age of onset (years)	Years of training	Instrument/s	Highest degree obtained
M1	5	48	Piano, recorder, hackbrett	Post graduate
M2	7	28	Piano, flute	Post graduate
M3	4	18	Piano, guitar, vocals, drums, bass guitar	Graduate
M4	7	13	Piano	Post graduate
M5	9	16	Guitar	Graduate
M6	4	22	Piano, violin	Doctorate
M7	9	20	Vocals, guitar	Post graduate
M8	5	20	Piano, vocals (soprano)	Post graduate
M9	5	16	Piano, alto saxophone, flute, vocals	Graduate
M10	7	17	Piano, vocals, guitar, trombone	Post graduate
M11	8	25	Piano, vocals	Doctorate
M12	5	19	Guitar	Post graduate
M13	6	10	Piano, vocals	Graduate
M14	7	15	Piano	Graduate
M15	9	12	Piano, violin	Graduate
M16	7	16	Violin, bass guitar	Graduate
M17	6	52	Piano, percussion	Doctorate
M18	4	25	Piano, bongo, percussion	Graduate
NM1	–	0	–	Post graduate
NM2	6	2	Piano	Doctorate
NM3	11	2	Trombone	Graduate
NM4	–	0	–	Graduate
NM5	–	0	–	Graduate
NM6	–	0	–	Graduate
NM7	10	3	Piano	Doctorate
NM8	–	0	–	Post graduate
NM9	–	0	–	Post graduate
NM10	–	0	–	Graduate
NM11	–	0	–	Graduate
NM12	–	0	–	Graduate
NM13	–	0	–	Doctorate
NM14	–	0	–	Graduate
NM15	–	0	–	Graduate
NM16	–	0	–	Graduate
NM17	–	0	–	Graduate
NM18	–	0	–	Post graduate
NM19	–	0	–	Post graduate
NM20	–	0	–	Graduate
NM21	–	0	–	Graduate
NM22	–	0	–	Graduate

Table 2
Details of various tests used with a brief description of procedure.

Measures	Tests	Procedure
Auditory processing	Frequency discrimination test (1 kHz) (Peter et al., 2014)	*Stimuli:* 1 kHz pure tone served as standard stimulus. The variable (target) stimuli were generated with frequencies ranging from 1001 Hz to 1050 Hz in steps of 1 Hz. All stimuli were 500 ms duration with a ramp of 20 ms. *Procedure:* Thresholds were estimated based on a 3 AFC procedure with a 2-down 1-up tracking method, estimating the 70.7% correct point on the psychometric function (Levitt, 1971). The target signal frequency was reduced after 2 correct responses and was increased after 1 incorrect response. *Response and scoring:* The participant's task was to identify the interval containing the different signal. The step size was initially 5 Hz and was reduced to 1 Hz after two reversals. The arithmetic mean of the last three reversals in a block of 6 was taken as threshold. Log transformation was applied to the thresholds.
	Threshold for discrimination of iterated rippled noise (Peter et al., 2014)	*Stimuli:* The iterated ripple noise (IRN) stimuli were generated using MATLAB 7 using the add-original configuration (IRNO) method described by Yost (1996). The standard stimulus was white noise with zero iteration. The variable IRNO stimuli were created by adding 10 ms delayed copies of white noise with the original noise. The process was repeated 8 times. IRNO were generated at different gain factors (g) that is, attenuation of the delayed repetition relative to the original noise in order to obtain versions of the stimuli with different pitch strengths. The g ranged from 0 to 0.2 in steps of 0.01. All the stimuli were 500 ms in duration with 30 ms rise and fall times. *Procedure:* Thresholds were estimated based on a 3 AFC procedure with a 2-down 1-up tracking method, estimating the 70.7% correct point on psychometric function (Levitt, 1971). In this procedure, the g of the target signal was reduced after 2 correct responses, and was increased after 1 incorrect response. *Response and scoring:* Participants indicated which of the three stimuli had a pitch percept. The step size for the variable stimuli was initially 0.02 and was reduced to 0.01 after two reversals. The arithmetic mean of the last six reversals in a block of 12 was taken as threshold.
	Gaps in noise test (Baker et al., 2008)	*Stimuli:* White noise with duration of 500 ms with ramp of 20 ms was used as the standard stimulus. White noise of 500 ms duration with varying durations of silence inserted in the centre were used as variable stimuli. *Procedure:* Thresholds were estimated based on a 3 AFC procedure with a 2-down 1-up tracking method, estimating the 70.7% correct point on psychometric function (Levitt, 1971). In this procedure, duration of silence in the target signal was reduced after 2 correct responses, and was increased after 1 incorrect response. *Response and scoring:* Participants indicated which of the three stimuli contained a gap. The step size was initially 3 ms and was reduced to 1 ms after two reversals. The arithmetic mean of the last 3 reversals in a block of 6 was taken as threshold.
	Detection of amplitude modulation (Peter et al., 2014)	*Stimuli:* The standard stimulus was a white noise low pass filtered at 20,000 Hz. The standard stimulus was amplitude modulated with varying modulation depths to create variable stimuli. The modulation frequencies used were 4 and 64 Hz. *Procedure:* 3 AFC with a 2-down 1-up tracking method, estimating the 70.7% correct point on psychometric function (Levitt, 1971) was employed to estimate threshold. In this procedure, the depth of modulation in the target signal was reduced after 2 correct responses, and was increased after 1 incorrect response. *Response and scoring:* Participants were asked to indicate which of the three stimuli was not constant. The amplitude modulation thresholds were based on the modulation depth in decibels ($20 \times \log_{10}(m)$). The step size was initially 4 dB and was reduced to 2 dB after two reversals. The arithmetic mean of the last three reversals in a block of 6 was taken as threshold.
	Dichotic digits test (3 pairs) (Strouse and Wilson, 1999)	*Stimuli:* Over the course of 25 trials, three pairs of digits were presented dichotically (total of 6 digits per trial). *Response and scoring:* After each trial, participants were asked to repeat as many of the 6 digits as they could. They were not asked to report digits from each ear separately. The recalled digits were scored according to which ear they were presented and percentage correct scores were obtained. The right ear advantage was calculated by subtracting left ear scores from right ear scores.
	Listening in Spatialised noise (sentences test) (Cameron and Dillon, 2007)	*Stimuli:* The commercially available Listening in Spatialized Noise Sentence test (LiSN-S) was used. Target sentences were mixed with distracter discourse and presented at a simulated 0° azimuth. Two subtests were administered: a) target and distractor discourse were spoken by the same speaker i.e. same voice-0° and b) target and distractor were spoken by different speakers i.e. different voices-0°. *Response and scoring:* Participants were asked to repeat target sentences and ignore the distractor discourse. The level of the target sentence was adjusted adaptively by the software to estimate the speech reception threshold (SRT) for the two subtests. The levels were adjusted in 4 dB steps until the first reversal in performance was recorded and in 2 dB steps thereafter.
Cognition	Forwards and backwards digit span test (Semel et al., 2006)	*Stimuli:* Digits were presented through headphones at the rate of one digit per second. *Response and scoring:* Participants were asked to recall the digits in same order (forward span) or reverse order (backward span). Age referenced scaled scores were obtained from raw scores using procedures described in CELF-IV (Australian) norms.
	Stroop Colour Word test	*Stimuli:* Computerised test which consisted of three subtasks: naming the ink colour in which the symbol 'X' is printed (Stroop I), reading the colour names printed in black ink (Stroop II), and naming the ink colour of the printed words in which ink color and the word differ (e.g., the word 'red' printed in green ink, Stroop III). The test was implemented in Presentation software (www.neurobs.com). *Response and scoring:* Four response buttons with names of colours printed in black ink were used. Participants pressed a button depending on whether they were asked to identify the ink colour or colour name. Reaction times were obtained for each trial. Mean reaction times were calculated for all the subtasks. Mean colour word interference score was calculated by subtracting the average time needed to complete the first two subtasks from the time needed to complete the third subtask (interference score = Stroop III - [(Stroop I + Stroop II)/2]) (Van der Elst et al., 2006).
	Integrated visual auditory continuous performance test (IVA-CPT) (Turner and Sandford, 1995)	*Stimuli:* Computerised test where 500 trials of visual and auditory '1's and '2's were presented pseudorandomly requiring a shift between the two modalities. *Response and scoring:* Over a span of 13 min, the participants were asked to click the mouse every time they saw or heard '1'. The number '2' was to be ignored. Sustained Auditory and Visual attention quotients are automatically generated by the reporting component of the IVA-CPT software, and represent age- and gender-matched population norms.

2.2.3.1. Stimuli used for familiarization.

The stimuli used for creating embedded triplets for the aSL familiarization phase were same as those described in previous studies (Abla et al., 2008; Saffran et al., 1999). Eleven pure tones within the same octave (starting at middle C or 261.6 Hz within a chromatic set) were generated using MATLAB (R2013 a). Tones are labelled according to their musical notation and are depicted in Fig. 1. The tones were 550 ms in duration (25 ms rise time and 25 ms fall time). Three tones were combined in succession to form a triplet. All participants were exposed to a familiarization stream containing 6 triplets (ADB, DFE, GG#A, FCF#, D#ED, CC#D) similar to those described by Saffran et al. (1999).

For the vSL task, stimuli comprised the 11 cartoons (described as 'aliens'), used by Arciuli and Simpson (2011, 2012a). Stimuli are depicted in Fig. 1. Cartoons were scaled such that the maximum height and width of all cartoons were equal. Each cartoon was presented for 550 ms (similar to tones in the aSL task) against a black background. Three cartoon figures were combined in succession to form a triplet. As with the aSL task, all participants were exposed to 6 visual triplets during familiarization.

2.2.3.2. Creation of familiarization stream.

Irrespective of the modality of the SL task, the familiarization streams were constructed in the following manner. Triplets were concatenated in a pseudorandom order to form a continuous stream of stimuli with no gaps between triplets. Thus, triplets were 'embedded' in a continuous stream. This stream consisted of 40 repetitions of each triplet. The familiarization stream was made up of 240 triplets (40 presentations of each of the 6 triplets) and was about 7 min long. Three familiarization streams were created with different pseudorandom ordering of triplets. The triplets were concatenated with two randomization constraints as described in previous studies (Arciuli and Simpson, 2011, 2012b): a) no repeated triplets were allowed and b) no repeated triplet pairs were allowed. The statistical structure for the aSL and vSL familiarization streams was identical. Statistical cues were the only indicator of triplet boundaries as there were no discernible discontinuities at the triplet boundaries.

Statistical cues to the presence of triplets included transitional probabilities (TP) within triplets and across triplet boundaries (Fig. 2). For example, the within-triplet TP for CC#D in Fig. 2 was calculated as the mean of TPs of the doublets 'CC#' and 'C#D'. For the aSL task and the vSL task, the TPs within triplets for all three streams ranged from 0.25 to 1 (mean 0.625). In contrast, the TPs across triplet boundaries were 0.04–0.3 (mean 0.11).

In the aSL familiarization stream, the mean frequencies for the initial, middle, and final tones across all triplets were 341 Hz (SD = 66), 321 Hz (SD = 57), and 370 Hz (SD = 82), respectively. There was no significant difference between the frequencies of the tones at each position within the tone triplets [$F(2, 10) = 1.48$, $p = 0.28$]. The mean frequencies of the tones within a triplet were as follows- ADB = 409.2 Hz, DFE = 324.2 Hz, GG#A = 415.8 Hz, FCF# = 326.9 Hz, D#ED = 311.5 Hz, and CC#D = 277.5 Hz. There was no significant difference between mean pitch interval within versus across triplets (3.1 vs 4.9 half tones in average).

2.2.3.3. Creation of test phase.

For each of the SL tasks, the test phase included 36 two alternative forced choice trials (2 AFC) (Abla and Okanoya, 2009; Saffran et al., 1999). For each trial during the test phase, participants were presented with an embedded triplet and a novel triplet (counterbalanced order of presentation). The novel triplets were made up from the same individual stimuli but had never occurred together as an embedded triplet in the familiarization stream. The novel triplets were formed according to those reported by Saffran et al. (1999). Each individual tone or cartoon within a triplet was presented using the same presentation time (550 ms) as had been used during familiarization. For each trial, the presentation of the embedded versus the novel triplet was separated by a 1-second gap. After the presentation of both triplets, a new screen appeared which prompted participants to identify which of the two triplets had appeared previously (during familiarization).

Before beginning the test phase, four practice trials were performed to ensure that participants waited for presentation of both triplets before responding. The practice trial triplets did not occur during the familiarization or test phases. This was done to avoid interference between the practice and test trials. Verbal feedback was given during the practice trials. No time constraints were imposed during the practice or test trials.

2.3. Procedure

Data were collected over two sessions. Hearing screening, auditory processing tests, and tests of cognition were administered in session 1. The aSL and vSL tests were administered in session 2. Stimuli for assessments in session 1 were presented using a laptop PC through headphones (Telephonics TDH-39) at 50 dB HL via an Interacoustics AC-40 clinical audiometer. The aSL and vSL task stimuli were delivered through Presentation software (Version 16.5, www.neurobs.com). The aSL stimuli were presented via Etymotic ER 3A insert ear phones (at 70 dB SPL). The visual stimuli were presented on a 17-inch CRT monitor placed at 1 m distance from the participant.

2.3.1. Familiarization phase for SL

Participants were exposed to auditory and visual stimulus streams during which they performed an oddball detection task. This task served as a cover task to ensure attentiveness. During the familiarization phase of the aSL task, participants were asked to press a button whenever they heard a pure tone with a frequency of 1319 Hz (not used in any embedded or novel triplets). During the vSL familiarization phase, they were asked to press a button whenever they saw a particular alien figure (not used in any embedded or novel triplets). To ensure that learning was implicit, participants were not given any instructions about the nature of the embedded triplets within the familiarization stream and were not told to learn or remember anything. There were 40 presentations of the oddball stimulus within each familiarization stream for each modality. The oddball stimulus was randomly presented at the end of a triplet.

Half the participants undertook the familiarization phase of the aSL task before the familiarization phase of the vSL task and the order was reversed for the other participants. The entire familiarization phase for both aSL and vSL lasted about 50 min (2 modalities X 3 streams X 7 min). Breaks of up to 5 min were given between presentations of the streams. Participants were informed about the upcoming test phase only after completion of the familiarization phases for both aSL and vSL tasks.

2.3.2. Test phase

The aSL and vSL test phases were administered in the same order as the familiarization phases. For instance, a participant who was first exposed to the aSL familiarization phase completed the aSL test phase task first. The duration of the test phase for each SL task was 5 min.

2.4. Data analysis

The data from 1 musician and 4 non-musicians were excluded for reasons including active middle ear pathology (1 non-musician)

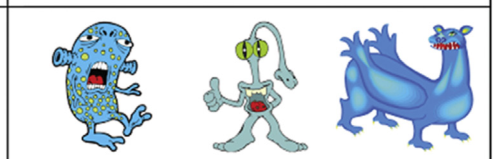
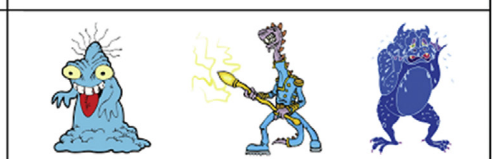
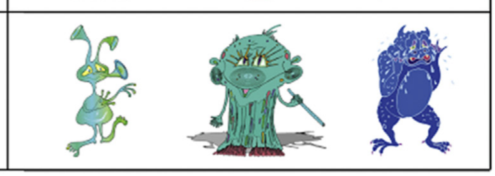
Triplet	aSL	vSL
Triplet 1	ADB	
Triplet 2	DFE	
Triplet 3	GG#A	
Triplet 4	FCF#	
Triplet 5	D#ED	
Triplet 6	CC#D	

Fig. 1. Triplets used for the aSL and vSL tasks. Stimuli were presented unimodally.

and a score of 0 on the IVA-CPT indicative of attention deficits. Hereafter, any description of the results does not include data for these 5 participants. Thus, data from 17 musicians (4 males, median age 26 years) and 18 non-musicians (4 males, median age 25.5 years) are reported. A Mann-Whitney *U* test showed no significant difference between the ages of the retained musicians and non-musicians [$U = 143.5$, $p = 0.76$].

Mann-Whitney U tests were used to compare performance on tests of auditory and cognitive processing between the two groups as the data were not normally distributed. Due to multiple comparisons for tests of auditory processing and cognition, we considered p values < 0.01 as significant. A Wilcoxon signed-rank test was used to assess if musical expertise was associated with modality specific enhancements in the attention tasks. Normality for performance on the SL tasks was confirmed using Shapiro-Wilk tests. Pearson's correlations were used to explore associations between SL and tasks of auditory or cognitive processing. All statistical analyses were performed using SPSS version 20 software.

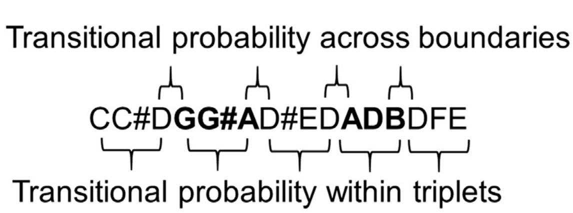


Fig. 2. Calculation of transitional probabilities (TP) for a representational aSL familiarization stream.

3. Results

3.1. Auditory processing

The mean, median, standard deviation, and results from the Mann-Whitney *U* test for all tests of auditory processing are shown in Table 3. A significant group difference was found for only one auditory processing measure: frequency discrimination. Musicians had better (lower) frequency discrimination thresholds than non-musicians.

3.2. Cognition

Table 4 presents the mean, median, standard deviation and statistics for all the measures of cognition. A significant group difference was found only for backwards digit span, with musicians exhibiting significantly higher scores than non-musicians. There were no group differences on tests of inhibition (Stroop task) and sustained attention. A Wilcoxon signed-rank test showed that the

Table 3
Means, medians, SDs and effect sizes for performance on tests of auditory processing. Significant results are in bold font.

Test	Musicians			Non-musicians			Statistics		
	Mean	Median	SD	Mean	Median	SD	U	*p*	r
Frequency discrimination test (log transformed)	**0.7**	**0.7**	**0.2**	**1.1**	**1**	**0.8**	**53.5**	**0.001**	**−0.6**
Threshold for discrimination of iterated rippled noise (IRNO gain factor)	0.07	0.05	0.02	0.07	0.07	0.03	123.5	0.3	−0.2
Gaps in noise test (ms)	2.7	2.6	0.4	2.6	2.6	0.4	133.5	0.5	−0.1
Detection of amplitude modulation (4 Hz)- Modulation depth (20 $\log_{10}$ m)	−23.7	−22.7	2.5	−23.3	−24	4.9	149.5	0.9	−0.01
Detection of amplitude modulation (64 Hz)- Modulation depth (20 $\log_{10}$ m)	−14.5	−14.7	3.3	−13.6	−12.7	2.9	121.5	0.3	−0.2
Dichotic digits test – right ear scores (%)	94.8	94.7	4.3	94.2	94.7	4.2	143.5	0.8	−0.4
Dichotic digits test – left ear scores (%)	94.4	94.7	5.3	89.9	89.3	6.5	82	0.02	−0.05
Dichotic digits test (right ear advantage)	0.4	0	5.9	4.4	4.7	5.1	94.5	0.05	−0.3
Listening in Spatialized noise different voices-0° (SRT in dB)	−7.8	−7.4	3.5	−7.9	−7.7	3.1	144.5	0.8	−0.04
Listening in Spatialized noise same voices-0° (SRT in dB)	−2.9	−2.5	2.2	−3.1	−2.5	2.2	139.5	0.7	−0.07

two groups performed at a similar level for auditory and visual attention tasks (musicians: $Z = -0.55$, $p = 0.6$; non-musicians: $Z = -1.45$, $p = 0.15$). There was no difference between musicians and non-musicians on either auditory or visual attention tasks. Additionally, modality specific enhancements in attention were not observed for musicians.

3.3. Statistical learning

Data from the oddball detection tasks during familiarization were analysed to determine the percentage of successfully identified stimuli. Each participant scored above 80% in the aSL and vSL oddball detection cover tasks. The percentage of correctly identified embedded triplets during the test phase was recorded for musicians and non-musicians. Consistent with previous SL studies (Arciuli and Simpson, 2012a; Conway et al., 2010; Stevens et al., 2015), participants who scored outside the range mean ± 2 SD were excluded from further analyses. This resulted in the exclusion of 4 participants – one non-musician for an excessively low score on the aSL task, one musician for an excessively high score on the vSL, and two non-musicians for excessively low scores on the vSL task. Thus, 34 participants were retained for aSL (17 musicians) and 32 participants for vSL (16 musicians).

Fig. 3 depicts the percentage of correctly identified embedded triplets for retained participants on the aSL and vSL tasks. We performed item analyses to ensure that responding was consistent across the six embedded triplets presented during the test phase. For example, the transitional probabilities were higher within some triplets than others. Results showed that responding was consistent across triplets for both groups. Normal distributions of performance on the aSL and vSL tasks was confirmed for both groups using the Shapiro-Wilk test (all $ps > 0.05$). One sample t-tests revealed that musicians and non-musicians performed significantly above chance on the aSL task [musicians $t(16) = 14.9$, $p < 0.001$, $d = 3.62$; non-musicians $t(16) = 6.1$, $p < 0.01$, $d = 1.49$]. Likewise, performance on the vSL task was significantly above chance for both groups [musicians $t(15) = 2.7$, $p < 0.05$, $d = 0.67$; non-musicians $t(15) = 2.9$, $p < 0.05$, $d = 0.73$].

We conducted a 2 × 2 ANOVA to compare performance on the two SL tasks (modality: aSL and vSL) across the two groups (musicians and non-musicians). There was a significant main effect of modality [$F(1, 29) = 61.57$, $p < 0.005$, partial $\eta^2 = 0.68$], and of group [$F(1, 29) = 5.71$, $p < 0.05$, partial $\eta^2 = 0.16$], and a significant interaction between group and modality [$F(1, 29) = 9.87$, $p < 0.01$, partial $\eta^2 = 0.25$]. Pairwise comparisons using t-tests with Bonferroni correction revealed that musicians significantly outperformed non-musicians on the aSL task [$t(30) = 3.29$, $p < 0.05$, $d = 1.13$] but not on the vSL task [$t(30) = -0.49$, $p > 0.05$, $d = 0.17$].

Scores on the aSL and vSL tasks were not correlated (all participants: $r = -0.02$, $p > 0.05$; musicians: $r = 0.37$, $p > 0.05$; non-musicians: $r = -0.15$. $p > 0.05$). For the musicians, there were no significant associations between SL performance and age of onset of music training (aSL: $r = 0.47$, $p > 0.05$; vSL: $r = 0.34$, $p > 0.05$) or between SL and years of music training (aSL: $r = -0.23$, $p > 0.05$; vSL: $r = -0.1$, $p > 0.05$).

3.4. Relationship between statistical learning and measures of auditory processing and cognition

Combining the data for both groups, we performed a Pearson's correlational analysis to explore the relationships between SL and measures of auditory processing and cognition. Scores for aSL were not correlated with any auditory processing or cognitive measures except frequency discrimination. There was a moderate negative correlation between performance on aSL and frequency discrimination thresholds ($r = -0.43$, $p = 0.01$). To tease apart the influence of musical expertise in this association, Pearson's r was calculated separately for each group. The results revealed no significant relationships (musicians: $r = 0.35$, $p = 0.17$; non-musicians: $r = -0.45$, $p = 0.07$).[1] vSL was not correlated with any of the measures of auditory processing or cognition. Based on the results of these correlational analyses and to avoid the risk of over-fitting, a follow up multiple-regression analysis was not conducted.

4. Discussion

Our findings indicate that musicians performed better than non-musicians in one auditory and one cognitive task: frequency discrimination and backward digit span. A key finding was that musicians outperformed non-musicians in identifying embedded triplets in the aSL task but not the vSL task. Performance on the SL tasks was not correlated with performance on the auditory and cognitive processing tasks.

4.1. Auditory processing

The musicians had lower frequency discrimination thresholds than the non-musicians and this difference was statistically significant. Enhanced frequency discrimination performance of musicians has been documented previously (Micheyl et al., 2006; Spiegel and Watson, 1984). Similar to the current study, Micheyl et al. (2006) found that musicians with classical music training of more than 10 years had lower frequency discrimination thresholds than non-musicians. Kishon-Rabin et al. (2001) found that classical

[1] We also confirmed this using a non-parametric Spearman's correlation. Performance on aSL and frequency discrimination was not correlated for either group (musicians: $r_s = 0.35$, $p = 0.17$; non-musicians: $r_s = -0.4$, $p = 0.11$).

Table 4
Means, medians, SDs and effect sizes for performance on tests of cognition. Significant results are in bold font.

Test	Musicians			Non-musicians			Statistics		
	Mean	Median	SD	Mean	Median	SD	U	*p*	r
Forwards digit span scaled scores	10.9	11	1.8	9.5	10	2.8	110.5	0.2	−0.24
Backwards digit span scaled scores	**12.2**	**13**	**2.5**	**9.5**	**9.5**	**2.1**	**57**	**0.001**	**−0.53**
Stroop colour word interference score (ms)	134.3	110.8	80.7	154.5	104	163.9	145	0.8	−0.04
Sustained auditory attention quotient	108.8	110	18.3	103.4	110	30.1	147.5	0.9	−0.03
Sustained visual attention quotient	106.1	111	18.0	96.9	105.5	28.7	105	0.1	−0.27

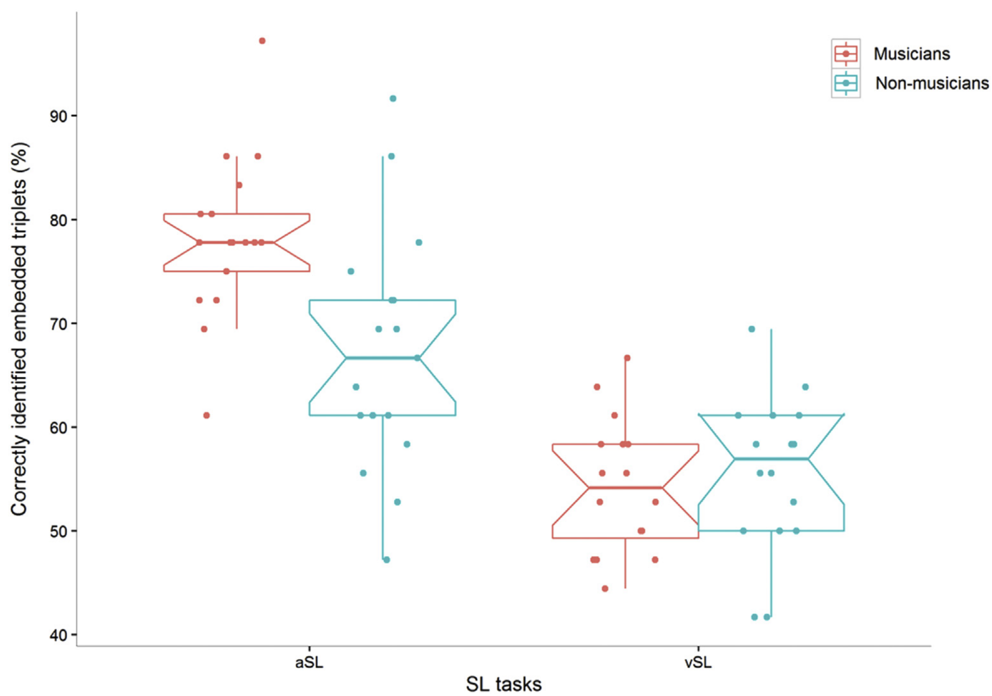


Fig. 3. Box and whisker plots showing the percentage of embedded triplets correctly identified by musicians and non-musicians on the aSL and vSL tasks. The box denotes 75th to 25th percentile (interquartile range). The upper whisker represents 75th percentile +1.5 × IQR and the lower whisker represents 25th percentile −1.5 × IQR. 50% indicates chance performance.

musicians had smaller frequency discrimination thresholds than contemporary (e.g. jazz, modern) musicians. It is noteworthy that the musicians in our study were also trained in classical music. Thus, smaller frequency discrimination thresholds may be due to emphasis on correct tuning during classical training (Micheyl et al., 2006).

We measured frequency discrimination at 1 kHz which has been suggested to be dominated by the use of temporal information that is, a phase locking mechanism (Moore, 2012). Musicians have been shown to have more precise temporal phase locking at the brainstem level as indexed by the frequency following response (FFR) (Lee et al., 2009). Better phase locking could be one of the reasons for finer frequency discrimination in musicians. Furthermore, it has been suggested that enhanced frequency discrimination could be due to better short-term memory representation or increased attention (Kishon-Rabin et al., 2001; Tervaniemi et al., 2005). However, for the musician group in our study, there was no correlation between frequency discrimination and performance on working memory and attention tasks. It is possible that the relationship between frequency discrimination and cognitive measures may be better probed by using different tests of attention and working memory than those used in the current study.

A right ear advantage (REA) was observed for both groups in the dichotic digits task. The difference in REA scores between the groups was not statistically significant. Musicians and non-musicians also had similar performance on the gap detection test. These findings are in agreement with previous literature comparing musicians and non-musicians on dichotic listening tasks involving digits (Nelson et al., 2003), and gaps in noise tasks (Ishii et al., 2006; Monteiro et al., 2010). Taken together, these findings suggest that superior performance of musicians may be limited to specific auditory tasks.

There were no group differences for detection of sinusoidal amplitude modulation of noise and discrimination of iterated rippled noise. Though a previous study (Lee et al., 2009) reported stronger encoding of temporal envelope cues in the brainstem of musicians using an electrophysiological technique (FFR), we found no evidence of this using behavioural measures. Enhancements in electrophysiological measures (such as increase in the amplitude of FFR) might not necessarily translate to enhancements in behavioural measures (Bidelman et al., 2011). Further research using more comparable behavioural and electrophysiological paradigms is needed to provide information about temporal envelope processing in musicians.

Musicians and non-musicians had similar speech perception in noise scores measured through the LiSN-S test. This is in line with three studies that also investigated speech perception in noise using a variety of tests (e.g. Boebinger et al., 2015; Fuller et al., 2014;

Ruggles et al., 2014). In contrast, in other studies of speech perception, musicians outperformed non-musicians (Parbery-Clark et al., 2009, 2011, 2012). These differences could be due to the different tests used across the studies. However, Ruggles et al. (2014), using the same tests and participant criteria as used by Parbery-Clark et al. (2009), found no significant advantage of being a musician on speech perception in noise tests (Experiment 2). Additionally, findings from Parbery-Clark et al. (2009) suggested that group differences may be observed only when the masking tasks are difficult, for example when the target and maskers are co-located. We administered two subtests of the LISN-S test, where the target speech and masker were co-located (different voices-0° and same voice-0°). However, there was no difference between the groups even when maskers were co-located and had no fundamental frequency cues (the hardest condition, i.e. same voice-0°). A recent study also found no significant group differences when the masker and target speech were co-located (Clayton et al., 2016). Further research involving stricter criteria for selection of participants (for example, recruitment of only professional musicians) and assessment of performance on a wider range of speech in noise tests may assist in understanding the relationship between speech perception in noise and musical experience. Overall, our results suggest that musical expertise might be associated with enhancements in only a subset of auditory perceptual skills.

4.2. Cognition

The current findings add to the converging evidence for musicians' superior performance on working memory tasks (Chan et al., 1998; Franklin et al., 2008; Ho et al., 2003). It has been suggested that musicians allocate more brain resources as the working memory load increases. This was evidenced by larger blood oxygenation-level dependent (BOLD) signal measured with functional magnetic resonance imaging (fMRI) in musicians compared to non-musicians during an n-back working memory task (Pallesen et al., 2010).

Interestingly, we found that musicians had significantly better scores on backward digit span but not on forward digit span. It has been suggested that distinct cognitive processes are tapped by forward and backward digit span tests and this may be why we observed that musicians outperformed non-musicians on backward but not forward digit span. For instance, the difference in performance on the forward and backward digit span tests may be explained by two theoretical approaches – the complexity view and the representational view (Rosen and Engle, 1997).

According to the complexity view, backward recall involves considerable attentional demands, with manipulation of digits held in short-term memory, and is thus considered to be a part of executive function processes tapping into working memory (Rosen and Engle, 1997). The second approach, the representational view, argues that backward recall may involve specific visuospatial processing where items are represented in a spatial array for easier reversal (Li and Lewandowsky, 1995; St Clair-Thompson and Allen, 2013). During the course of training and practice, musicians are required to learn melody and memorize sequences of notes either by learning through ear or through visual memory of music notations. A reasonable hypothesis to explain these group differences is that musicians improve their working memory storage through practice (consistent with the complexity view). Alternatively, musicians may employ visual strategies (such as visualizing each digit) to assist them with backward recall (consistent with the representational view). Anecdotally, a few musicians in our study reported using visual strategies during the backward recall task.

In the current study, musicians did not have enhanced performance on the visual Stroop task which is consistent with evidence that musical experience may be linked with specific components of executive function such as backwards recall but not inhibition (Boebinger et al., 2015; Clayton et al., 2016; Zuk et al., 2014). It is noteworthy that Bialystok and DePape (2009) using an *auditory* Stroop task demonstrated enhanced inhibition in musicians. Further studies are required to understand the influences of modality on executive function tasks.

Although it has been suggested that musical training contributes to enhanced auditory but not visual attention (Strait et al., 2010), our results indicated no enhancements in either visual or auditory sustained attention. This could be attributed to the different tasks and types of attention measured in the two studies. Strait and colleagues measured *alertness* in one modality at a time by comparing reaction times in the presence or absence of a variable delay cue. In contrast, our study used the IVA-CPT test (Table 2) which measured *sustained* attention in auditory and visual modalities concurrently by accounting for accuracy as well as reaction time in the presence of a distracting stimulus. Overall, the mixed pattern of results observed between this and other studies suggests that the relationship between musical expertise and general cognitive abilities is complex and needs further investigation.

4.3. Statistical learning

To the best of our knowledge, this is the first study showing behavioural differences in aSL between musicians and non-musicians using an embedded triplet task comprising tones. In previous studies that have assessed aSL using both electrophysiological and behavioural measures in musicians and non-musicians, group differences were observed only for electrophysiological measures (François and Schön, 2011; Paraskevopoulos et al., 2012). However, it should be noted that behavioural test results in the latter study did not differ significantly from chance for either group. In our study, both groups showed a mean aSL that was significantly better than chance, however, musicians outperformed non-musicians.

Though the current research design cannot address the question of whether musical training causes enhancement of aSL in adults, a randomized longitudinal training study of 8-year old children suggested a causal relationship between musical training and aSL (François et al., 2013). The enhanced aSL for musicians in our study may be attributed to the fact that Western tonal music is based on a strong system of regularities between musical events, such as notes and chords (Tillmann et al., 2000). Our findings suggest that long-term musical expertise could help form associations between successive stimuli and group them into distinctive units (triplets in our case) purely based on the transitional probabilities. Although the triplets we used did not follow the rules of any standard music composition, it could be argued that musicians' pre-existing knowledge regarding relative pitches of the Western scale could make it easier to detect statistical regularities amongst tones using the Western scale (Loui, 2012). Using a new, unfamiliar Bohlen-Pierce scale, Loui and colleagues reported no differences in performance of musically trained and untrained subjects for melody recognition and rule generalization (Loui et al., 2010). However, it should be noted that the musician group in that study had less musical training than the participants in our study (range of 5–14 years with an average of 9.6 years). In contrast, the musicians in our study had at least 10 years' experience of singing or playing music (range of 10–52 years with an average of 21.6 years). A systematic study using professional musicians or musicians with at least 10 years of musical training might help to clarify if the duration of the musical training is relevant in statistical learning of unfamiliar musical scales.

Notwithstanding the fact that some previous research has shown that musical expertise is associated with an enhanced visuospatial iconic representation (Gromko and Poorman, 1998), and visual imagery (Neuhoff et al., 2002), musicians and non-musicians in the current study had comparable performance on the vSL task. This finding could be taken as support for an experience-dependent modality specific enhancement of SL that is, the musicians' advantage in SL was only seen in one modality (aSL) and not the other modality (vSL). It is possible that more efficient usage of visuospatial perception and imagery in musicians (Brochard et al., 2004) might translate into enhancements in vSL measured using stimuli that contains spatial regularities. Further research is needed to explore these possibilities.

Better performance in the aSL task than in the vSL task by both groups suggests that SL proceeds differently across the visual and auditory modalities (Conway and Christiansen, 2005, 2009). Additionally, there was no significant correlation of the scores for the aSL and vSL tasks. This is in line with recent research (Frost et al., 2015; Siegelman and Frost, 2015). Given that musical practice involves the integration of multi-modal stimuli, future studies may further explore the role of musical expertise in SL using an auditory-visual SL task.

4.4. Relationship between statistical learning and measures of auditory processing and cognition

Shook et al. (2013) postulated that enhanced SL in musicians may be due to their enhanced temporal processing of auditory information as reflected in temporal discrimination, gap detection and rhythm perception tasks (Rammsayer and Altenmüller, 2006). However, in our study, performance on both SL tasks was independent of performance on any auditory processing task for both musicians and non-musicians. Future studies could use an aSL task involving stimuli with closer frequency separation to advance our knowledge about the relationship between aSL and frequency discrimination in musicians. In addition, further investigation using a variety of auditory processing tasks involving music perception (such as chord perception, melody discrimination) could help shed light on what strategies musicians use for extracting distributional cues.

Our finding that SL performance was not related to cognitive ability is in line with previous studies where SL was not associated with measures of verbal reasoning, intelligence and working memory (Kaufman et al., 2010). Siegelman and Frost (2015) also reported that there was no relationship between scores on SL (aSL and vSL tasks using the embedded triplet paradigm) and a large battery of cognitive tests (working memory, verbal working memory, rapid automatized naming, and switch task). These findings point towards the independence of SL from other general cognitive abilities. However, it should be noted that the relationship between SL and cognitive abilities such as working memory is complex and warrants further investigation. For instance, Arciuli and Simpson (2011) argued that implicit rather than explicit working memory tasks might reveal a relationship with SL. Furthermore, SL performance measured through 2AFC trials may involve additional processes such as decision making (François et al., 2014). Future studies may investigate the relationship between SL and cognitive abilities by incorporating concurrent cognitive tasks as SL takes place and using a wide battery of SL tasks (including SL tasks that do not rely on 2AFC trials).

It has been hypothesized that SL operates automatically with little or no dependence on executive attentional resources (Turk-Browne et al., 2005). Indeed, our results showed that SL performance was not correlated with measures of attention and inhibition. For both SL tasks used in our study, the participants were engaged in a cover task during familiarization that required them to process the stimuli in a different way (to detect oddballs rather than to extract statistical information per se). It is notable that the participants were successful in recognising embedded triplets in both the aSL and vSL tasks despite this competing demand that meant they were actively trying to process the stimuli in another way. As long as participants attend to the relevant stimuli, aSL and vSL may indeed operate automatically.

As music consists of a variety of temporal patterns, future research may investigate the link between aSL and other aspects of auditory processing, such as rhythm perception. Finally, a longitudinal study involving music training and measurement of SL, along with a comprehensive auditory processing test battery could shed some light on causality and the mechanisms underlying enhanced aSL.

4.5. Summary

In summary, using a set of auditory processing, cognitive, and statistical learning measures, we found that long-term musical expertise is associated with better performance on frequency discrimination and backward digit span tasks. We also present the first empirical evidence of better aSL, but not vSL for musicians than for non-musicians as evaluated using the embedded triplet paradigm. Performance on SL tasks was not correlated with performance on any other auditory processing or cognition measures.

Acknowledgements

We would like to thank the editor Brian Moore and three anonymous reviewers for their helpful comments. This work was funded by the Australia Awards Scholarships and the HEARing CRC, established and supported under the Cooperative Research Centres Program – Business Australia. This work was also supported by the Australian Research Council Centre of Excellence for Cognition and its Disorders (CE110001021). Demuth was supported by a Laureate Fellowship granted by the Australian Research Council (FL130100014). Arciuli was supported by a Future Fellowship granted by the Australian Research Council (FT130101570).

References

Abla, D., Katahira, K., Okanoya, K., 2008. On-line assessment of statistical learning by event-related potentials. J. Cognitive Neurosci. 20 (6), 952–964.
Abla, D., Okanoya, K., 2009. Visual statistical learning of shape sequences: an ERP study. Neurosci. Res. 64 (2), 185–190. http://dx.doi.org/10.1016/j.neures.2009.02.013.
Arciuli, J., Simpson, I.C., 2011. Statistical learning in typically developing children: the role of age and speed of stimulus presentation. Dev. Sci. 14 (3), 464–473. http://dx.doi.org/10.1111/j.1467-7687.2009.00937.x.
Arciuli, J., Simpson, I.C., 2012a. Statistical learning is lasting and consistent over time. Neurosci. Lett. 517 (2), 133–135. http://dx.doi.org/10.1016/j.neulet.2012.04.045.
Arciuli, J., Simpson, I.C., 2012b. Statistical learning is related to reading ability in children and adults. Cognitive Sci. 36 (2), 286–304. http://dx.doi.org/10.1111/j.1551-6709.2011.01200.x.
Arciuli, J., von Koss Torkildsen, J., 2012. Advancing our understanding of the link between statistical learning and language acquisition: the need for longitudinal data. Front. Psychol. 3, 324. http://dx.doi.org/10.3389/fpsyg.2012.00324.
Baker, R.J., Jayewardene, D., Sayle, C., Saeed, S., 2008. Failure to find asymmetry in auditory gap detection. Laterality 13 (1), 1–21.
Bialystok, E., DePape, A.-M., 2009. Musical expertise, bilingualism, and executive functioning. J. Exp. Psychol. Hum. Percept. Perform. 35 (2), 565–574.
Bidelman, G.M., Gandour, J.T., Krishnan, A., 2011. Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. Brain Cognition 77 (1), 1–10.
Boebinger, D., Evans, S., Rosen, S., Lima, C.F., Manly, T., Scott, S.K., 2015. Musicians and non-musicians are equally adept at perceiving masked speech. J. Acoust. Soc. Am. 137 (1), 378–387. http://dx.doi.org/10.1121/1.4904537.
Brandler, S., Rammsayer, T.H., 2003. Differences in mental abilities between musicians and non-musicians. Psychol. Music 31 (2), 123–138.
Brochard, R., Dufour, A., Després, O., 2004. Effect of musical expertise on visuospatial abilities: evidence from reaction times and mental imagery. Brain

Cognition 54 (2), 103–109. http://dx.doi.org/10.1016/S0278-2626(03)00264-1.
Cameron, S., Dillon, H., 2007. Development of the listening in spatialized noise-sentences test (LISN-S). Ear Hear. 28 (2), 196–211.
Chan, A.S., Ho, Y.-C., Cheung, M.-C., 1998. Music training improves verbal memory. Nature 396 (6707), 128–128.
Clayton, K.K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A.D., Kidd Jr., G., 2016. Executive function, visual attention and the cocktail party problem in musicians and non-musicians. PloS one 11 (7), e0157638. http://dx.doi.org/10.1371/journal.pone.0157638.
Conway, C.M., Bauernschmidt, A., Huang, S.S., Pisoni, D.B., 2010. Implicit statistical learning in language processing: word predictability is the key. Cognition Int. J. Cognitive Sci. 114 (3), 356–371. http://dx.doi.org/10.1016/j.cognition.2009.10.009.
Conway, C.M., Christiansen, M.H., 2005. Modality-constrained statistical learning of tactile, visual, and auditory sequences. J. Exp. Psychol. Learn. Mem. Cognition 31 (1), 24.
Conway, C.M., Christiansen, M.H., 2009. Seeing and hearing in space and time: effects of modality and presentation rate on implicit statistical learning. Eur. J. Cognitive Psychol. 21 (4), 561–580.
Fine, P.A., Moore, B.C., 1993. Frequency analysis and musical ability. Music Percept. An Interdiscip. J. 11 (1), 39–53.
François, C., Chobert, J., Besson, M., Schön, D., 2013. Music training for the development of speech segmentation. Cereb. Cortex 23 (9), 2038–2043.
François, C., Jaillet, F., Takerkart, S., Schön, D., 2014. Faster sound stream segmentation in musicians than in nonmusicians. PloS one 9 (7), e101340.
François, C., Schön, D., 2010. Learning of musical and linguistic structures: comparing event-related potentials and behavior. Neuroreport 21 (14), 928–932. http://dx.doi.org/10.1097/WNR.0b013e32833ddd5e.
François, C., Schön, D., 2011. Musical expertise boosts implicit learning of both musical and linguistic structures. Cereb. Cortex 21 (10), 2357–2365. http://dx.doi.org/10.1093/cercor/bhr022.
François, C., Schön, D., 2014. Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: the role of musical practice. Hear. Res. 308, 122–128.
Franklin, M.S., Moore, K.S., Yip, C.-Y., Jonides, J., Rattray, K., Moher, J., 2008. The effects of musical training on verbal memory. Psychol. Music 36 (3), 353–365.
Frost, R., Armstrong, B.C., Siegelman, N., Christiansen, M.H., 2015. Domain generality versus modality specificity: the paradox of statistical learning. Trends Cognitive Sci. 19 (3), 117–125.
Fuller, C.D., Galvin, J.J., Maat, B., Free, R.H., Başkent, D., 2014. The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? Front. Neurosci. 8, 179. http://dx.doi.org/10.3389/fnins.2014.00179.
Gaser, C., Schlaug, G., 2003. Brain structures differ between musicians and non-musicians. J. Neurosci. 23 (27), 9240–9245.
Gromko, J.E., Poorman, A.S., 1998. Does perceptual-motor performance enhance perception of patterned art music? Music. Sci. 2 (2), 157–170.
Helmbold, N., Rammsayer, T., Altenmüller, E., 2005. Differences in primary mental abilities between musicians and nonmusicians. J. Individ. Differ. 26 (2), 74–85.
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 (3), 439–450.
Iliadou, V.V., Bamiou, D.E., Chermak, G.D., Nimatoudis, I., 2014. Comparison of two tests of auditory temporal resolution in children with central auditory processing disorder, adults with psychosis, and adult professional musicians. Int. J. Audiol. 53 (8), 507–513. http://dx.doi.org/10.3109/14992027.2014.900576.
Ishii, C., Arashiro, P.M., Pereira, L.D., 2006. Ordering and temporal resolution in professional singers and in well tuned and out of tune amateur singers. Pró-Fono Rev. Atualização Científica 18 (3), 285–292.
Jakobson, L.S., Lewycky, S.T., Kilgour, A.R., Stoesz, B.M., 2008. Memory for verbal and visual material in highly trained musicians. Music Percept. An Interdiscip. J. 26 (1), 41–55.
Kaufman, S.B., DeYoung, C.G., Gray, J.R., Jiménez, L., Brown, J., Mackintosh, N., 2010. Implicit learning as an ability. Cognition Int. J. Cognitive Sci. 116 (3), 321–340. http://dx.doi.org/10.1016/j.cognition.2010.05.011.
Kishon-Rabin, L., Amir, O., Vexler, Y., Zaltz, Y., 2001. Pitch discrimination: are professional musicians better than non-musicians? J. Basic Clin. Physiol. Pharmacol. 12 (2), 125–144.
Lee, K.M., Skoe, E., Kraus, N., Ashley, R., 2009. Selective subcortical enhancement of musical intervals in musicians. J. Neurosci. 29 (18), 5832–5840.
Lee, Y.-s., Lu, M.-j, Ko, H.-p., 2007. Effects of skill training on working memory capacity. Learn. Instr. 17 (3), 336–344.
Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49 (2B), 467–477.
Li, S.-C., Lewandowsky, S., 1995. Forward and backward recall: different retrieval processes. J. Exp. Psychol. Learn. Mem. Cognition 21 (4), 837.
Loui, P., 2012. Statistical learning – What can music tell us?us. In: Rebuschat, P., Williams, J. (Eds.), Statistical Learning and Language Acquisition. Mouton de Gruyter, Berlin, Germany, pp. 433–462.
Loui, P., Wessel, D.L., Kam, C.L.H., 2010. Humans rapidly learn grammatical structure in a new musical scale. Music Percept. An Interdiscip. J. 27 (5), 377–388.
Micheyl, C., Delhommeau, K., Perrot, X., Oxenham, A.J., 2006. Influence of musical and psychoacoustical training on pitch discrimination. Hear. Res. 219 (1), 36–47.
Monteiro, R.A.M., Nascimento, F.M., Soares, C.D., Ferreira, M.I.D.d.C., 2010. Temporal resolution abilities in musicians and no musicians violinists. Arq. Int. Otorrinolaringol. 14 (3), 302–308.
Moore, B.C.J., 2012. An Introduction to the Psychology of Hearing. Brill, Netherlands.
Münte, T.F., Altenmüller, E., Jäncke, L., 2002. The musician's brain as a model of neuroplasticity. Nat. Rev. Neurosci. 3 (6), 473–478.
Nelson, D.M., Wilson, R.H., Kornhass, S., 2003. Performance of musicians and nonmusicians on dichotic chords, dichotic CVs, and dichotic digits. J. Am. Acad. Audiol. 14 (10), 536–544.
Neuhoff, J.G., Knight, R., Wayand, J., 2002. Pitch change, sonification, and musical expertise: Which way is up?. In: Paper Presented at the International Conference on Auditory Display.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113.
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 (6), e11120.
Paraskevopoulos, E., Kuchenbuch, A., Herholz, S.C., Pantev, C., 2012. Statistical learning effects in musicians and non-musicians: an MEG study. Neuropsychologia 50 (2), 341–349. http://dx.doi.org/10.1016/j.neuropsychologia.2011.12.007.
Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. Ear Hear. 30 (6), 653–661.
Parbery-Clark, A., Strait, D.L., Kraus, N., 2011. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. Neuropsychologia 49 (12), 3338–3345.
Parbery-Clark, A., Tierney, A., Strait, D.L., Kraus, N., 2012. Musicians have fine-tuned neural distinction of speech syllables. Neuroscience 219, 111–119.
Patel, A.D., 2010. Music, Language, and the Brain. Oxford university press.
Patel, A.D., 2011. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. Front. Psychol. 2, 142. http://dx.doi.org/10.3389/fpsyg.2011.00142.
Peter, V., Wong, K., Narne, V.K., Sharma, M., Purdy, S.C., McMahon, C., 2014. Assessing spectral and temporal processing in children and adults using temporal modulation transfer function (TMTF), iterated ripple noise (IRN) perception, and spectral ripple discrimination (SRD). J. Am. Acad. Audiol. 25 (2), 1–9.
Rammsayer, T., Altenmüller, E., 2006. Temporal information processing in musicians and nonmusicians. Music Percept. An Interdiscip. J. 24 (1), 37–48.
Rodrigues, A.C., Guerra, L.B., Loureiro, M.A., 2007. Visual attention in musicians and non-musicians: a comparative study. In: Paper Presented at the Proceedings of the 3rd International Conference on Interdisciplinary Musicology.
Rohrmeier, M., Rebuschat, P., 2012. Implicit learning and acquisition of music. Top. Cognitive Sci. 4 (4), 525–553.
Rosen, V.M., Engle, R.W., 1997. Forward and backward serial recall. Intelligence 25 (1), 37–47.
Ruggles, D.R., Freyman, R.L., Oxenham, A.J., 2014. Influence of musical training on understanding voiced and whispered speech in noise. PloS one 9 (1), e86980.
Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-Month-old infants. Science 274 (5294), 1926–1928.
Saffran, J.R., Johnson, E.K., Aslin, R.N., Newport, E.L., 1999. Statistical learning of tone sequences by human infants and adults. Cognition Int. J. Cognitive Sci. 70 (1), 27–52.
Schlaug, G., Norton, A., Overy, K., Winner, E., 2005. Effects of music training on the child's brain and cognitive development. Ann. N. Y. Acad. Sci. 1060 (1), 219–230.
Schön, D., François, C., 2011. Musical expertise and statistical learning of musical and linguistic structures. Front. Psychol. 2, 167. http://dx.doi.org/10.3389/fpsyg.2011.00167.
Semel, E.M., Wiig, E.H., Secord, W., 2006. CELF 4: Clinical Evaluation of Language Fundamentals. Pearson: Psychological Corporation.
Shook, A., Marian, V., Bartolotti, J., Schroeder, S.R., 2013. Musical experience influences statistical learning of a novel language. Am. J. Psychol. 126 (1), 95.
Siegelman, N., Frost, R., 2015. Statistical learning as an individual ability: theoretical perspectives and empirical evidence. J. Mem. Lang. 81, 105–120.
Skoe, E., Krizman, J., Spitzer, E., Kraus, N., 2013. The auditory brainstem is a barometer of rapid auditory learning. Neuroscience 243, 104–114. http://dx.doi.org/10.1016/j.neuroscience.2013.03.009.
Špajdel, M., Jariabková, K., Riečanský, I., 2007. The influence of musical experience on lateralisation of auditory processing. Laterality 12 (6), 487–499.
Spiegel, M.F., Watson, C.S., 1984. Performance on frequency-discrimination tasks by musicians and nonmusicians. J. Acoust. Soc. Am. 76 (6), 1690–1695.
St Clair-Thompson, H.L., Allen, R.J., 2013. Are forward and backward recall the same? A dual-task study of digit recall. Mem. Cognition 41 (4), 519–532.
Stevens, D.J., Arciuli, J., Anderson, D.I., 2015. Concurrent movement impairs incidental but not intentional statistical learning. Cognitive Sci. 39 (5), 1081–1098.
Strait, D.L., Kraus, N., 2011. Playing music for a smarter ear: cognitive, perceptual and neurobiological evidence. Music Percept. An Interdiscip. J. 29 (2), 133.
Strait, D.L., Kraus, N., Parbery-Clark, A., Ashley, R., 2010. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. Hear. Res. 261 (1), 22–29.
Strouse, A., Wilson, R.H., 1999. Recognition of one-, two-, and three-pair dichotic digits under free and directed recall. J. Am. Acad. Audiol. 10 (10), 557–571.
Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., Schröger, E., 2005. Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. Exp. Brain Res. 161 (1), 1–10.
Tillmann, B., Bharucha, J.J., Bigand, E., 2000. Implicit learning of tonality: a self-organizing approach. Psychol. Rev. 107 (4), 885–913.
Turk-Browne, N.B., Junge, J., Scholl, B.J., 2005. The automaticity of visual statistical

learning. J. Exp. Psychol. General 134 (4), 552–564. http://dx.doi.org/10.1037/0096-3445.134.4.552.
Turner, A., Sandford, J., 1995. A normative study of IVA: Integrated visual and auditory continuous performance test. In: Paper Presented at the Annual Convention of the American Psychological Association.
Van der Elst, W., Van Boxtel, M.P., Van Breukelen, G.J., Jolles, J., 2006. The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. Assessment 13 (1), 62–79. http://dx.doi.org/10.1177/1073191105283427.
Yost, W.A., 1996. Pitch strength of iterated rippled noise. J. Acoust. Soc. Am. 100 (5), 3329–3335.
Zendel, B.R., Alain, C., 2012. Musicians experience less age-related decline in central auditory processing. Psychol. Aging 27 (2), 410. http://dx.doi.org/10.1037/a0024816.
Zuk, J., Benjamin, C., Kenyon, A., Gaab, N., 2014. Behavioral and neural correlates of executive functioning in musicians and non-musicians. PloS one 9 (6), e99868. http://dx.doi.org/10.1371/journal.pone.0099868.