Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *The Neurosciences and Music IV: Learning and Memory*

A sensitive period for musical training: contributions of age of onset and cognitive abilities

Jennifer Bailey and Virginia B. Penhune

Psychology Department, Concordia University, Montreal, Quebec, Canada

Address for correspondence: Jennifer Bailey, Concordia University – Psychology, 7141 Sherbrooke Street West, Montreal, Quebec H4R 1B6, Canada. j.anne.bailey@gmail.com

The experiences we engage in during childhood can stay with us well into our adult years. The idea of a sensitive period—a window during maturation when our brains are most influenced by behavior—has been proposed. Work from our laboratory has shown that early-trained musicians (ET) performed better on visual-motor and auditory-motor synchronization tasks than late-trained musicians (LT), even when matched for total musical experience. Although the groups of musicians showed no cognitive differences, working memory scores correlated with task performance. In this study, we have replicated these findings in a larger sample of musicians and included a group of highly educated nonmusicians (NM). Participants performed six woodblock rhythms of varying levels of metrical complexity and completed cognitive subtests measuring verbal abilities, working memory, and pattern recognition. Working memory scores correlated with task performance across all three groups. Interestingly, verbal abilities were stronger among the NM, while nonverbal abilities were stronger among musicians. These findings are discussed in context of the sensitive period hypothesis as well as the debate surrounding cognitive differences between musicians and NM.

Keywords: sensitive period; musicians; cognitive abilities

Introduction

The plastic changes that occur in each of our brains as we mature are the result of an interaction between maturational changes and experience. A fascinating example of this interaction is a "sensitive period"—a window of time during development when brain systems are more susceptible to the influence of experience or stimulation. In our lab, we have used trained musicians to study possible sensitive period effects. In these studies, musicians who began training before age 7 demonstrated enhanced rhythm synchronization performance compared to musicians who began their training later in development, when matched for total musical experience.[1,2] In addition, although these two groups of highly trained musicians did not differ on global cognitive variables, individual working memory scores predicted synchronization performance. In this study, the sample size has been increased and a nonmusician's (NM) group has been added to further elucidate the association between working memory, musical training, and task performance. Including a group of highly educated NM also provides insight into possible cognitive differences between musicians and NM.

As our knowledge about brain plasticity evolves, evidence for sensitive periods related to the acquisition of a variety of skills increases. The idea of a sensitive period may have gained most widespread attention through the results of a number of studies showing that second-language proficiency is greater in individuals who were exposed to the second language before age 11–13.[3,4] Recent evidence using neuroimaging techniques also supports the idea that the sensory systems have developmental windows of time during which they are most sensitive to stimulation. Differences in occipital recruitment for nonvisual functions between congenitally blind individuals and those who acquired blindness later in

doi: 10.1111/j.1749-6632.2011.06434.x
 © 2012 New York Academy of Sciences.

development suggest that the visual system also has a developmental window during which it is most responsive to stimulation.[5] Cochlear implantation studies suggest that the auditory system is more responsive the earlier these devices are implanted.[6,7] Studies have reported differences in brain structure between early-trained (ET) and late-trained (LT) musicians and have associated these differences with the extent of musical experience.[8–10] However, an important addition to the investigation of a sensitive period is the matching paradigm developed in our laboratory.[2] When ET and LT musicians are matched for musical experience (e.g., years of formal instruction, years of playing, current hours of practice), the general effects associated with musical experience are controlled and the age at which they began their musical training is isolated as the variable of interest.

Evidence from previous studies in our lab supports the idea of a sensitive period among musicians, even when cognitive abilities are considered. ET musicians (those who began before age 7) outperformed LT musicians (those who began after age 7) on an auditory–motor synchronization task, as well as a visual–motor synchronization task, when matched for total musical experience.[1,2] The two groups did not differ on cognitive measures such as vocabulary (VC), matrix reasoning (MR), digit-span (DS), and letter–number sequencing (LN).[1,11,12] However, working memory scores predicted performance on the rhythm synchronization task across both groups of musicians. A regression analysis revealed that after controlling for working memory, group membership still accounted for variance in task performance. These results suggest that a musician's working memory and age of starting musical training both contributed to their ability to perform the rhythm synchronization task.

This study aims to replicate our previous findings in a larger sample of musicians, and shed light on the debate surrounding cognitive differences between musicians and NM. Although cognitive differences between musicians and NM have been reported, there is controversy in the literature over how or why these differences emerge.[13,14] Studies have used child samples to examine the interaction between music lessons and cognitive and brain development.[15,16] Using an adult sample complements studies of children by allowing us to test whether differences associated with musical training persist into adulthood, especially because we are comparing musicians to a group of highly educated NM. In addition, using a group of adult musicians with extensive but variable lengths of musical training allows us to investigate the nature of the association between music lessons and cognitive abilities.

Method

Participants

Fifty neurologically healthy individuals between the ages of 18 and 36 ($M = 25.5$ years old, SD $= 4.6$) participated in this study. Participants were screened for significant head injuries, history of neurological disease, or medication that could affect task performance. Of the 50 participants, 30 were highly trained and currently practicing musicians and 20 were NM (less than three years of musical experience). The musical training and experience of each participant was determined through a Musical Experience Questionnaire (MEQ) that was developed within our laboratory.[1] The MEQ quantifies the amount of instrumental, vocal, and dance training an individual has received; at what age this training occurred; and the amount of time currently dedicated to practicing on a weekly basis. All musicians had extensive musical experience ($M = 16.4$ years; SD $= 4.4$). Musicians were classified as ET ($n = 15$) or LT ($n = 15$) musicians, based on their MEQ data. Those who began their musical experience prior to or at the age of 7 were placed in the ET group and those who began after the age of 7 were classified as LT. The age of 7 was chosen based on previous studies.[1,2,7] The two groups were matched on years of musical experience, years of formal training, and hours of current practice. All participants gave informed consent and the Concordia University Research Ethics Committee had approved the protocol.

Stimuli

The rhythm task used in this study consisted of six woodblock rhythms of varying difficulty based on their metrical structure.[17,18] Each test rhythm consisted of 11 woodblock notes and had a total duration of six seconds. These rhythms differed in their temporal structure, such that the inter onset intervals between musical notes varied, but not the duration of the notes themselves. More specifically, each rhythm was made up of five eighth notes (each 250 ms), three quarter notes (each 500 ms), one

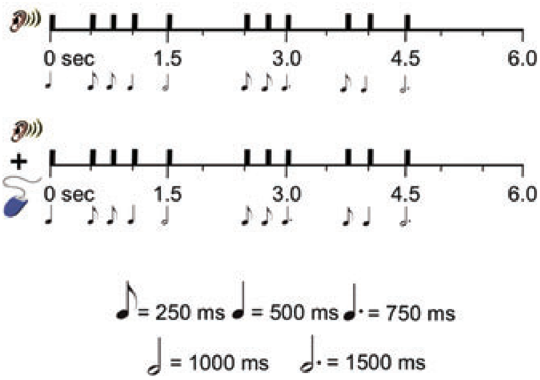

Figure 1. Illustration of the rhythm task. Participants were exposed to six rhythms presented in random order for approximately two 12-min blocks. Two different rhythms of each rhythmic complexity were used. Each trial consisted of a listening component followed by a listening and tapping component.

dotted quarter note (750 ms), one half note (1000 ms), and one dotted half note (1,500 ms). Manipulation of the temporal structure of the notes resulted in progressively more complex and less metrically structured rhythms. For a more detailed description of this task and the metrical complexity manipulation, please see Bailey and Penhune.[1]

Participants completed the DS and LN subtests from the Wechsler Adult Intelligence Scale–III (WAIS) and the VC and MR subtests from the Wechsler Abbreviated Scale of Intelligence (WASI).[10,11] The DS requires individuals to recall strings of numbers, and the LN requires individuals to recall and mentally manipulate strings of letters and numbers. Both of these subtests tap into working memory abilities; however, LN imposes a heavier load on working memory, while DS consists of a rote auditory memory recall section in addition to a mental manipulation section. The VC assesses an individual's ability to orally define words and the MR assesses nonverbal reasoning and visual pattern recognition abilities. Both VC and MR are strongly correlated with global IQ, although they assess different types of intelligence.

Procedure

During the rhythm task, participants alternated between listening and tapping along while each rhythm played twice (Fig. 1). Participants were instructed to tap as accurately as possible with the rhythm as it played during the tapping repetition. Two very basic practice rhythms were administered to familiarize participants with the task. Each rhythm presented in a counterbalanced fashion six times over approximately 12 min in each block and participants performed two blocks. Once participants had completed the first block of the task, they were asked to perform DS. Participants then performed the second block of the rhythm synchronization task, followed by VC, LN, and finally, MR.

Measures

Musical information was quantified for each participant in terms of years of experience, years of formal training, and hours of current weekly practice using the MEQ.[1] Individual cognitive abilities were measured using the four chosen cognitive subtests (DS, LN, VC, and MR). Results were scored according to standard procedure. Performance on the rhythm synchronization task was measured using three dependent variables: percent correct (PC), asynchrony (ASYN), and inter-tap interval (ITI) deviation. A tap was considered correct if it was made within half of

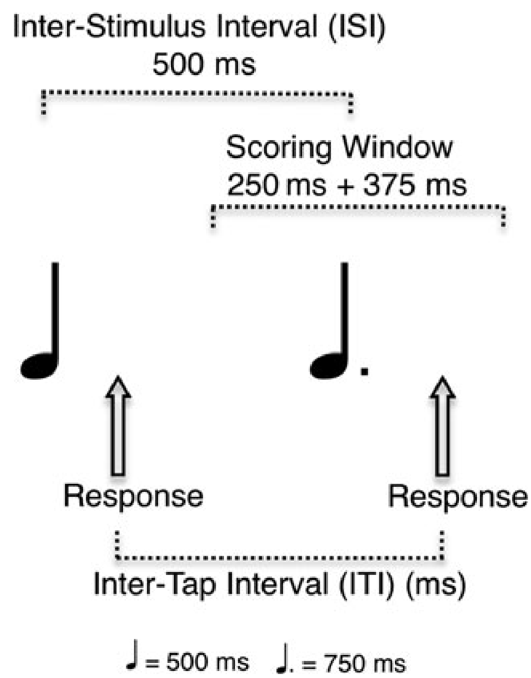

Figure 2. Illustration of the scoring method used to evaluate rhythm task performance. A response was scored correctly if the mouse tap was made within half of the onset-to-onset interval before and after a woodblock note. Asynchrony was measured as the difference between each woodblock note and the participant's response. ITI deviation was calculated as a ratio of the ITI and the ISI subtracted from 1.

the onset-to-onset interval before or after a woodblock note (Fig. 2). The ASYN measure was defined as the absolute value of temporal difference between the onset of each woodblock note and the associated mouse key press. The ITI deviation measure indicated the extent of deviation of the participant's tap interval from the actual interval between each pair of woodblock notes. It was calculated by dividing the interval between each pair of the participant's taps by the interval between each corresponding pair of woodblock notes in the rhythms and subtracting this ratio from a value of one. This measure is indicative of how well participants reproduced the temporal structure of the rhythms.

Data analysis

To compare rhythm synchronization across the three groups, a repeated-measures analysis of variance (ANOVA) for each of the dependent variables was conducted, with group as the between-subjects factor and rhythm type as the within-subjects factor. Pair-wise comparisons for between group differences were analyzed using least significant differences (LSD) correction for multiple comparisons. The result of our matching procedure was evaluated using t-test analyses for years of musical experience, years of formal training, and hours of current practice among the musicians. Group differences on the cognitive subtests were assessed using a one-way ANOVA for each cognitive variable with group as the between-subjects factor. Pair-wise comparisons were conducted using an LSD correction for multiple comparisons. The relationships among cognitive measures, musical experience variables, and task performance were examined using one-tailed Pearson correlation analyses. Raw scores on the cognitive subtests were used to correlate with performance measures and scaled scores were used when comparing the three groups on the cognitive measures. However, results were consistent whether raw or scaled scores were used in the analyses.

Based on a previously observed relationship between individual working memory abilities and task performance among musicians, a hierarchical regression analysis was conducted with all three groups in order to assess whether the observed group difference persists after individual working memory scores are considered.[1] A model was created with total ITI deviation as the dependent measure and both group and working memory as predictors. A composite score for each participant's working memory ability was created by summing their LN and DS scores and used in the regression analysis.

Results

Group comparisons of musical and cognitive measures between the ET and LT musicians confirmed that the two groups were well matched in terms of years of musical experience, years of formal training, and hours of current practice (Table 1). The one-way ANOVA revealed no significant differences in DS or LN scores between groups, although statistical trends toward a main effect of group on MR and VC were observed (Fig. 3). Pair-wise comparisons revealed that the NM VC scores were higher than the ET ($P = 0.026$), and the MR scores of the LT were higher than those of the NM ($P = 0.017$). Scaled scores were used for these analyses.

Behavioral measures

The ANOVA comparing accuracy (PC) of the rhythm reproduction task across the three groups showed a significant main effect of group ($F[2, 47] = 3.99$, $P < 0.05$; Fig. 4A). Pair-wise comparisons using an LSD correction revealed differences between the ET and NM ($P < 0.01$). These results confirm that all three groups were performing the

Table 1. Group demographics of musical experience variables

Group	Age (years)	Age of onset (years)	Formal training (years)	Musical experience (years)	Current practice (years)
Early-trained	23.47 (±3.85)	5.87 (±1.19)	11.73 (±3.97)	16.87 (±4.10)	15.23 (±9.97)
Late-trained	26.60 (±5.22)	10.47 (±2.03)	10.03 (±4.39)	15.90 (±4.74)	14.43 (±7.80)
t-values	–1.87	–7.57*	1.11	0.60	0.25

Note: Standard deviation values are in brackets.
*P value < 0.001.

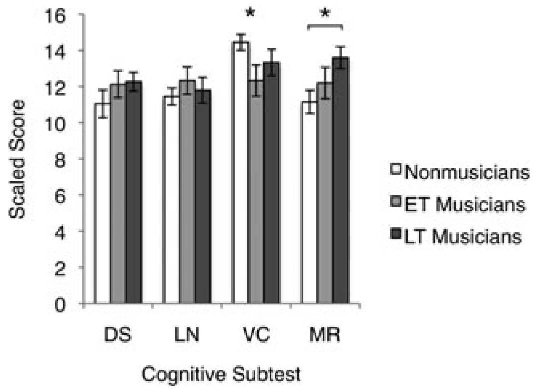

Figure 3. Group mean cognitive scaled scores. DS, digit-span; LN, letter–number sequencing; VC, vocabulary; MR, matrix reasoning. No group differences were observed on the two measures of working memory (DS and LN); however, statistical trends toward group differences were observed on VC ($P = 0.078$) and MR ($P = 0.055$). Pair-wise comparisons revealed specific group differences (*$P < 0.05$).

task correctly overall and the mean performance values were in the expected order (i.e., ET > LT > NM).

The ANOVA comparing performance on the synchronization measure (ASYN) across the three groups revealed a similar pattern of results, such that there was a main effect of group ($F\ [2, 47] = 16.76$, $P < 0.001$; Fig. 4B). Pair-wise comparisons using an LSD correction revealed lower ASYN scores for the ET and LT when compared to the NM ($P < 0.001$ for both comparisons). In addition, the ET group was better able to synchronize their responses than the LT musician group as revealed by lower ASYN scores ($P = 0.05$). These results suggest that the group differences were heightened on this more sensitive performance measure compared to our more global measure of accuracy (PC).

Consistent with the other performance measures, the ANOVA comparing reproduction of the temporal structure of the rhythms using our ITI measure of deviation across the three groups showed a significant main effect of group ($F[2, 47] = 20.30$, $P < 0.001$; Fig. 4C). Pair-wise comparisons using an LSD correction revealed a similar pattern of results as on the ASYN measure such: The ET had lower deviation scores than the LT ($P < 0.05$) and both musician groups had lower deviation scores than the NM ($P < 0.001$ for both comparisons).

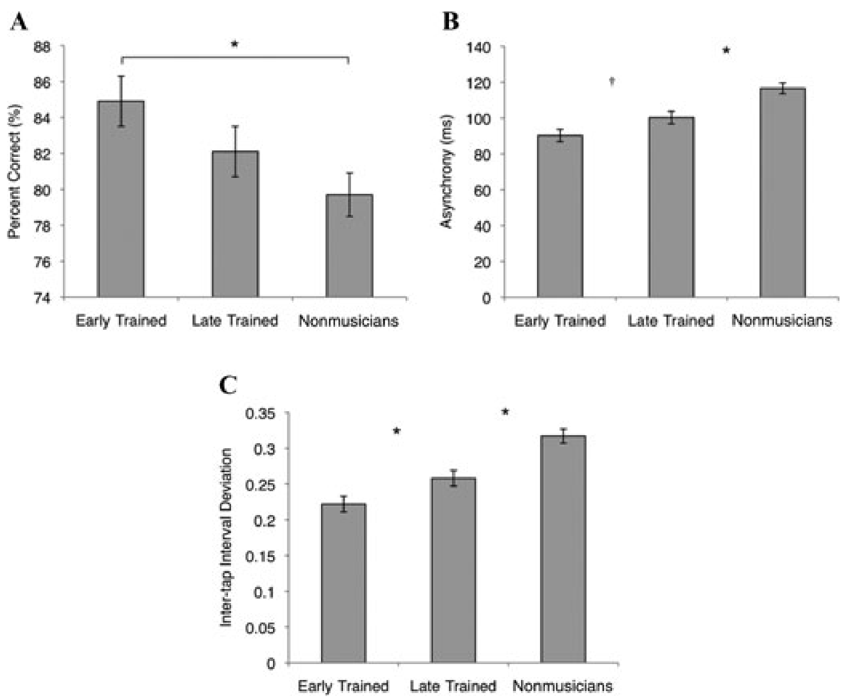

Figure 4. Task performance results for all three groups: (A) percent correct; (B) asynchronization; and (C) inter-tap interval deviation. Repeated measures ANOVA for each performance measure revealed a significant main effect of group, and pair-wise comparisons revealed specific group differences (*$P < 0.05$, † $= 0.05$). Standard error bars have been used.

Table 2. Correlation results between cognitive scores and task performance measures

Performance measure	Digit-span	Letter–number sequencing	Matrix reasoning	Vocabulary
Percent correct	0.275***	0.360*	0.147	–0.072
Asynchrony	–0.258***	–0.307*	–0.262***	0.269***
Inter-tap interval deviation	–0.378**	–0.340*	–0.339*	0.187

Note: Raw scores were used for the cognitive measures.
* P values < 0.05.
** P values < 0.01.
*** P values < 0.08 but greater than 0.05.

These results further illustrate that as the measure of performance increased in sensitivity to temporal aspects of the rhythms, the observed group differences were heightened.

Correlations

In order to examine the relationship between task performance and cognitive variables across the three groups, raw scores for PC, ASYN, and ITI were correlated with raw scores for VC, MR, DS, and LN (Table 2). Both working memory measures (DS and LN) correlated significantly with the three performance measures (PC, ASYN, and ITI) in the expected directions, confirming that the rhythm reproduction task implicates the use of working memory. Surprisingly, VC correlated with the synchronization measure such that higher VC scores were associated with poorer performance of the rhythm task. In addition, MR correlated with both synchronization and ITI measures of task performance. These results were likely driven by the group differences observed on these cognitive scores and this will be addressed in the discussion section.

Results from the correlational analyses between the behavioral measures and musical variables in the musicians (Table 3) reveal a significant association between years of formal training and ITI deviation ($r = –0.367$, $P < 0.05$). In addition, age of onset showed a significant relationship with ASYN and ITI, as well as a relationship trend toward significance with PC. In order to examine the association among years of formal training, cognitive scores, and task performance, correlations were performed between years of formal training and each cognitive measure. This set of analyses revealed an association trend toward significance between years of formal training and DS ($r = 0.342$, $P = 0.06$); however, no significant associations with LN, VC, or MR.

Regression analysis

In order to determine if the amount of variance in ITI deviation during task performance accounted for by group was above and beyond what was explained by individual working memory abilities, a hierarchical regression analysis was conducted using the three groups (Table 4). These results confirmed that, while individual working

Table 3. Correlation results between musical experience and task performance measures

Performance measure	Age of onset (years)	Formal training (years)	Musical experience (years)	Current practice (hours)
Percent correct	–0.352***	0.010	0.141	–0.052
Asynchrony	0.459*	–0.214	–0.139	–0.079
Inter-tap interval deviation	0.509**	–0.367*	–0.095	0.046

Note: Raw scores were used for the cognitive measures.
* P values < 0.05.
** P values < 0.01.
*** P values < 0.08 but greater than 0.05.

Table 4. **Regression analysis results examining the predictive value of group membership above and beyond working memory to task performance**

	R^2	β	R^2 change	F
Step 1	0.165		0.165	9.45
Working memory		–0.406*		
Step 2	0.538			27.336
Working memory		–0.293*		
group		0.621*	0.373	

Note: A working memory composite score was used for this analysis comprised of individual raw DS and LN scores.
* P values < 0.01.

memory abilities were predictive of task performance, group membership accounted for additional portions of the variance in ITI deviation scores.

Discussion

These findings replicate our previous findings but in a larger sample, and provide further evidence for a sensitive period for musical training that may have a specific impact on sensorimotor synchronization abilities. In this study, the ET musicians were better able to reproduce the rhythms than the LT musicians, even after controlling for years of formal training, playing experience, and current hours of practice. In addition, the two musician groups did not differ on the four cognitive measures. In other words, this observed group difference on task performance could not be attributed to differences in musical experience or cognitive ability, but to the developmental window during which musical training began. As expected, NM rhythm synchronization abilities were inferior to both musician groups. Although there were no differences in working memory performance across the three groups, individual working memory scores correlated with task performance, suggesting similar reliance on working memory resources for all groups. In further support of the sensitive period hypothesis, the regression results suggest that even after controlling for individual working memory scores, group membership still predicted a significant amount of variance in task performance. This reinforces the idea that musical training, especially early musical training, improves task performance above and beyond the contribution of working memory abilities.

In addition to the differences between ET and LT musicians on the rhythm task, we observed differences in performance on global cognitive variables between musicians and NM. Specifically, the NM obtained higher VC scores, while the musician groups had higher MR scores. These findings are interesting and can shed light on the types of cognitive effects associated with musical training. One hypothesis is that music lessons benefit the underlying cognitive abilities that are measured by MR, and that in contrast, NM are exercising their verbal intelligence via other avenues. If this were the case, one would expect the length of musical training (i.e., years of formal training) to be correlated with MR scores among the musician group, but it is not. Alternatively, one could hypothesize that individuals with strong visual-spatial organization skills are inclined to take up music lessons, and those with strong verbal abilities are likely to take up other NM activities. If this were true, then no relationship between length of musical training and MR would be present, yet group differences would persist between musicians and NM. The current data supports this assumption. The more general question of what is driving cognitive differences between musicians and NM is an area of controversy. Recently, Schellenberg and Peretz proposed that the observed association between music lessons and cognition may be mediated by executive function, although a more recent publication by Schellenberg failed to report convincing evidence that this was the case.[12,13] In our sample, a weak association between working memory, a component of executive function, and years of formal training was observed among the musicians; however, their scores were not higher than the NM, suggesting that if years of formal training impacts working memory, it does not do so above and beyond other nonmusical activities in which NM engage. Other factors, such as socioeconomic status or the family environment, may contribute to the differences between the two groups. Both our musicians and NM were either in the process of completing an undergraduate degree or had obtained one, and some were pursuing higher-level education. Thus, in these highly educated samples, any enhanced cognitive abilities observed in musicians over NM are likely to be a combination of innate predisposition and effects associated with exercising the abilities implicated in music lessons during development. Similarly, NM are likely predisposed to engage in other nonmusical activities and exercise other abilities during their development.

In summary, this study adds to the growing literature in support of a sensitive period for sensorimotor-integration abilities among musicians and considers NM as a comparison sample. Any differences in brain structure between early and LT musicians associated with these enhanced synchronization abilities have yet to be explored. The results from this study also add to the evidence that musicians and NM possess different cognitive strengths, even in a sample of highly educated adults. However, the exact contributions of innate predisposition and the influence of training remain unknown.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Bailey, J.A., & V.B. Penhune. 2010. Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. *Exp. Brain Res.* **204:** 91–101.
2. Watanabe, D., T. Savion-Lemieux & V. Penhune. 2007. The effect of early musical training on adult motor performance: evidence for a sensitive period in motor learning. *Exp. Brain Res.* **176:** 332–340.
3. Johnson, J.S. & E.L. Newport. 1989. Critical periods effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* **21:** 60–99.
4. Weber-Fox, C. & H.J. Neville. 2001. Sensitive periods differentiating processing of open- and closed-class words: an ERP study of bilinguals. *J. Speech Lang. Hear. Res.* **44:** 1338–1353.
5. Voss, P., F. Gougoux, R. Zatorre, *et al*. 2008. Differential occipital responses in early- and late-blind individuals during a sound-source discrimination task. *NeuroImage* **40:** 746–758.
6. Kral, A., R. Hartmann, J. Tillein, *et al*. 2001. Delayed maturation and sensitive periods in the auditory cortex. *Audio. Neuro-Otol.* **6:** 346–362.
7. Sharma, A., P. Gilley, M. Dorman & R. Baldwin. 2007. Deprivation-induced cortical reorganization in children with cochlear implants. *Int. J. Audiol.* **46:** 494–499.
8. Schlaug, G., L. Jäncke, Y. Huang, *et al*. 1995. Increased corpus callosum size in musicians. *Neuropsychologia* **33:** 1047–1055.
9. Bengtsson, S.L., Z. Nagy, S. Skare, *et al*. 2005. Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* **8:** 1148–1150.
10. Imfeld, A., M. Oechslin, M. Meyer, *et al*. 2009. White matter plasticity in the corticospinal tract of musicians: a diffusion tensor imaging study. *NeuroImage* **46:** 600–607.
11. Wechsler, D. 1997. *Wechsler Adult Intelligence Scale*. 3rd ed. Psychological Corporation. San Antonio, TX.
12. Wechsler, D. 1999. *Wechsler Abbreviated Scale of Intelligence*. Psychological Corporation. San Antonio, TX.
13. Schellenberg, E. & I. Peretz. 2008. Music, language and cognition: unresolved issues. *Trends Cogn. Sci.* **12:** 45–46.
14. Schellenberg, E.G. 2011. Examining the association between music lessons and intelligence. *Br. J. Psychol. (London, England: 1953)* **102:** 283–302.
15. Hyde, K.L., J. Lerch, A. Norton, *et al*. 2009. Musical training shapes structural brain development. *J. Neurosci.* **29:** 3019–3025.
16. Schellenberg, E.G. 2006. Long-term positive associations between music lessons and IQ. *J. Educ. Psychol.* **98:** 457–468.
17. Essens, P. 1995. Structuring temporal sequences: comparison of models and factors of complexity. *Percep. Psychophys.* **57:** 519–532.
18. Essens, P. & D. Povel. 1985. Metrical and nonmetrical representations of temporal patterns. *Percep. Psychophys.* **37:** 1–7.