Music lessons, pitch processing, and g

Psychology of Music

Psychology of Music
Copyright © 2009
Society for Education, Music
and Psychology Research
209–221
10.1177/0305735609339473
http://pom.sagepub.com

E. GLENN SCHELLENBERG AND SYLVAIN MORENO
UNIVERSITY OF TORONTO, CANADA

ABSTRACT Musically trained and untrained participants were administered tests of pitch processing and general intelligence (*g*). Trained participants exhibited superior performance on tests of pitch-processing speed and relative pitch. They were also better at frequency discrimination with tones at 400 Hz but not with very high tones (4000 Hz). The two groups also performed similarly on a measure of *g*. The findings suggest that music training is associated positively with various aspects of pitch processing for tones in the typical pitch range for music. They also imply that general associations between music lessons and nonmusical cognitive functioning stem from individual differences in psychological mechanisms distinct from *g*.

KEYWORDS: *executive function, intelligence, music lessons, pitch perception, relative pitch*

Introduction

Scholarly interest in associations between music lessons and cognitive abilities has grown in recent years. Comparisons of musically trained and untrained participants represent natural experiments that have ramifications for issues central to cognitive science, including plasticity (Trainor, 2005), modularity (Peretz & Coltheart, 2003), talent (Howe, Davidson, & Sloboda, 1998), and transfer (Schellenberg, 2005, 2006a). Here we report results from a natural experiment that investigated whether music lessons are associated with cognitive and perceptual abilities.

Music lessons are known to be associated positively with pitch processing, including melody recognition (Orsmond & Miller, 1999) and noticing whether a sequence of chords ends in a manner typical of Western music (e.g. Koelsch, Jentschke, Sammler, & Mietchen, 2007). Such advantages extend to tasks that involve the perception of pitch in speech (Magne, Schön, & Besson, 2006; Marques, Moreno, Castro, & Besson, 2007; Moreno et al., 2009; Schön, Magne, & Besson, 2004; Thompson, Schellenberg, & Husain, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). To test the limits of associations between music lessons and pitch processing, we administered five different tasks. One involved pitch-processing speed, two involved relative pitch, and two involved frequency discrimination – one with stimuli in the typical pitch range for music, the other with much higher stimuli.

We also administered a test of general intelligence (*g*). Although musicians do *not* consistently exhibit cognitive advantages over nonmusicians (Brandler & Rammsayer, 2003; Helmbold, Rammsayer, & Altenmüller, 2005), natural experiments reveal that music lessons in childhood have positive associations with mathematical, spatial, and verbal abilities in childhood and adulthood (for reviews see Schellenberg, 2005, 2006a). One possibility is that higher IQs increase the likelihood of taking music lessons and of performing well on a variety of cognitive measures. There is some evidence, however, that the causal direction goes *from* music lessons *to* cognitive functioning, at least in part (Moreno et al., 2009; Schellenberg, 2004). For example, when six-year-olds are assigned randomly to a year of music lessons (keyboard or vocal) or to control conditions (drama or no lessons), the music groups show larger increases in full-scale IQ (FSIQ) over the course of first grade (Schellenberg, 2004). As in the natural experiments, the advantage is general, extending across the various subtests and indexes of the Wechsler Intelligence Scale for Children – Third Edition (WISC-III, Wechsler, 1991).

In correlational studies or natural experiments, associations between music lessons and IQ could be due to confounding variables such as socioeconomic status or parents' education. Nonetheless, when large samples of children and undergraduates are administered the WISC-III and the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III, Wechsler, 1997), respectively, duration of music lessons is associated positively with FSIQ, even when family income, parents' education, and involvement in other extracurricular activities are held constant (Schellenberg, 2006b). Again, such associations are evident across the components of IQ tests but strongest for the aggregate measures (FSIQ and the principal component). On the one hand, then, the available literature implies that music lessons enhance *g*. On the other hand, approximately 50 percent of *g* is genetic in origin, whereas environmental influences remain poorly understood (e.g. Deary, Spinath, & Bates, 2006; Petrill et al., 2004; Plomin & Spinath, 2004).

In the present study, we tried to resolve this apparent conundrum by administering Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998) to musically trained and untrained adults. Raven's test is considered to be the best stand-alone measure of *g* (e.g. Carpenter, Just, & Shell, 1990). In two previous studies (Thompson et al., 2004; Trimmer & Cuddy, 2008), musically trained undergraduates outperformed their untrained counterparts on Raven's test. Nonetheless, both studies used a considerably shortened version of the test (Bors & Stokes, 1998), which made it a 'quite different task' (Hamel & Schmittmann, 2006, p. 1045).

The pitch-processing tasks served as control measures for Raven's test and vice versa. If the musically trained and untrained groups differed on Raven's test *and* on one or more aspects of pitch processing, analyses of covariance could determine whether between-group differences on one measure were independent of differences on other measures. By contrast, if the two groups did not differ on Raven's test, differences in pitch processing would confirm that the sample size was large enough to detect predicted group differences. We expected that music lessons would be predictive of better performance on the pitch-processing tasks except for the high-frequency discrimination task, which comprised tones outside the typical pitch range for music. We doubted, however, that associations between music lessons and cognitive abilities would extend to *g* as measured by Raven's test.

Method

PARTICIPANTS

The participants were 40 undergraduates (30 female, 10 male; mean age of 19.9 years, *SD* = 1.6 years) registered in an introductory psychology course who were recruited on the basis of their musical background. Half had extensive training in music, which included at least 8 years of lessons. Each year of lessons on two or more instruments was considered as two years, and each additional year of playing regularly (beyond lessons) was considered to be equivalent to half a year of lessons. Using these criteria, the group had 14.3 years of lessons on average (*SD* = 3.8). Each member had played regularly up until 3 years (or fewer) before participating (*M* = 0.7 years, *SD* = 0.9). Primary instruments were typically piano (*n* = 13) or violin (*n* = 4) but many participants had studied more than one instrument (*M* = 2.0, *SD* = 1.1). The other 20 participants had little or no musical training (same criteria, *M* = 1.6 years, *SD* = 2.0). Gender and musical training were counterbalanced. Participants received partial course credit plus US$10 for participating in the experiment, which consisted of two test sessions that lasted approximately 2.5 hours in total.

APPARATUS

The pitch-processing tasks were run on an iMac computer while participants sat in a sound-attenuating booth. Software created with PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) was used to present stimuli and collect responses. Stimuli were created using SoundEdit software and saved as digital sound files.

STIMULI AND MEASURES

Pitch-processing speed

Pitch-processing speed was measured with *auditory inspection time* (Deary, 1995); a task that does not require a speeded response. On each trial, listeners heard two pure tones (sine waves) that differed in pitch by 2 semitones, a difference that even five-year-olds with no music lessons can judge as 'higher' or 'lower' (Stalinski, Schellenberg, & Trehub, 2008). The task was to identify the order of the tones (high–low or low–high). The higher tone was 880 Hz (A_5, or one octave above concert A) and the lower tone was 784 Hz (G_5). Both test tones were well within the typical pitch range for music. Each trial began with a warning tone of 500 ms (832 Hz), followed by 1s of silence, followed by the two test tones with order determined randomly on each trial. To mask echoic memory, the test tones were followed by a warble: the same two tones alternating rapidly (5 ms each) for 1 s.

A three-up, one-down adaptive procedure was used to determine a threshold for each listener, specifically the briefest duration at which they performed well (85 percent correct). After three correct responses at a given duration, the duration of the test tones became briefer. After one incorrect response, the duration became longer. After nine reversals in duration (from increases to decreases, or decreases to increases), the test session ended. Each listener's threshold was the average duration of the test tones on the last three reversals. Tone durations had 30 possible levels (see Table 1), varying from relatively long (1 s) to very short (15 ms). To prevent audible clicks, stimulus tones between 40 ms and 1 s had linear onset and offset ramps of 10 ms. Tones of briefer duration (15 to 30 ms) had ramps of 5 ms.

Frequency discrimination

Two tasks measured frequency discrimination. In the *low-frequency* task, listeners heard two pure tones of 300 ms (onset/offset ramps of 10 ms) presented in succession. One tone was a standard fixed at 400 Hz, approximately midway in pitch between middle C and C_5 (one octave higher). The other tone was always higher in pitch. On each trial, participants' task was to decide if the higher tone was presented first or second. The actual order was determined randomly. The tones were separated initially by an interval of 3 semitones. The pitch difference became smaller as the trials progressed and listeners responded correctly, using the adaptive procedure

TABLE 1 *Stimulus parameters for the adaptive procedures used in the pitch-processing tasks*

Level	Processing speed (ms)	Frequency discrimination (cents)	Relative pitch (cents)
1 (easiest)	1000	300	200
2	900	250	180
3	800	200	160
4	750	150	140
5	700	125	120
6	650	100	100
7	600	90	90
8	550	80	80
9	500	70	70
10	450	60	60
11	400	50	50
12	350	40	45
13	300	30	40
14	250	25	35
15	225	20	30
16	200	15	25
17	175	10	20
18	150	5	15
19	125	3	10
20	100	1	5
21	90		
22	80		
23	70		
24	60		
25	50		
26	40		
27	30		
28	25		
29	20		
30 (hardest)	15		

described above. The pitch difference had 20 possible levels, with the easiest (first) level having a large difference (3 semitones) and the most difficult level having a very small difference (1 cent; a cent is one hundredth of a semitone, see Table 1).

The high-frequency discrimination task was identical to the low-frequency task except that the standard had a frequency of 4000 Hz rather than 400 Hz. In contrast to the low-frequency task, stimulus tones were higher than the pitch of most musical tones. The standard was just slightly lower in pitch than C_8 (the highest note on a standard piano, 4 octaves above middle C).

Relative pitch

Two tasks tested memory for pitch relations. On one, listeners heard the first two lines of 'Twinkle Twinkle Little Star' (14 tones: 'Twin-kle twin-kle lit-tle star. How I won-der what you are'). Their task was to judge whether the melody was presented in-tune or out-of-tune. The melody comprised pure tones (onset/offset ramps of 10 ms) of two durations. The longer tones ('star' and 'are') were 1 s; the shorter tones (other syllables) were 500 ms. Mistunings involved downward pitch displacements to the *dominant* note of the scale (*sol*), which occurred three times in the melody (i.e. tones 3, 4, and 7). Using the same adaptive procedure described above, mistunings were largest on the initial trials but became increasingly smaller on consecutive trials if listeners responded correctly. There were 20 possible levels of mistuning, ranging from 2 semitones to 5 cents (see Table 1). Whether the melody was in-tune or out-of-tune was determined randomly on each trial. To avoid the use of absolute-pitch cues from level to level, the first tone was middle C at odd-numbered levels and four semitones higher (E_4) at even-numbered levels.

The other relative-pitch task was identical except that the melody comprised the first two lines of 'Happy Birthday To You' (12 tones: 'Hap-py birth-day to you. Hap-py birth-day to you') and mistunings involved upward displacements to the *tonic* (*do*), which occurred twice in the melody (tones 5 and 12). The most frequent tone duration (the syllables in 'birthday to') was 600 ms, the longest duration was 1.2 s ('you'), the first syllable of 'happy' was 400 ms, and the second syllable was 200 ms. For both relative-pitch tasks, component tones of the stimuli were well within the typical pitch range for music (between middle C and E_5).

Raven's test

General cognitive ability (*g*) was measured with Raven's Advanced Progressive Matrices (36 items; Raven et al., 1998). Each item consisted of a 3 by 3 matrix, with line drawings of abstract patterns in all but the bottom-right cell. Participants were required to select the missing pattern from eight possible alternatives. The items became progressively more difficult, such that they required greater reasoning ability and intellectual capacity. Raw scores (number correct) were used in the analyses.

PROCEDURE

For each pitch-processing task, participants wore headphones and recorded their judgments by clicking the mouse on one of two buttons presented on the monitor. Feedback was provided after each response. During the first of two test sessions, participants completed a musical background questionnaire. They were also familiarized with the adaptive procedure by completing the processing-speed task three times, with a short break between each run. The first run was a practice run.

At the second session (different day), each of the remaining four auditory tasks was completed twice, with a short break between runs. Half of the participants completed the frequency-discrimination tasks before the relative-pitch tasks. The other half had the relative-pitch tasks first. The low-frequency discrimination task preceded the high-frequency task for half of the participants, with the order reversed for the other half. Similarly, 'Twinkle Twinkle' preceded 'Happy Birthday' for half of the participants, with the order reversed for the other half. Thus, the frequency-discrimination and relative-pitch tasks were presented in four different orders, with order counterbalanced across the musically trained and untrained groups. Participants were subsequently allowed a maximum of 40 minutes to complete Raven's test in a quiet room.

Results

Performance was consistent across the two runs for each of the five pitch-processing tasks (processing speed: $r = .97$, low-frequency discrimination: $r = .64$, high-frequency discrimination: $r = .58$, relative pitch – 'Twinkle': $r = .94$, relative pitch – 'Happy': $r = .82$, Ns = 40, ps < .01). Subsequent analyses examined thresholds averaged across the two runs. Correlations among the six outcome measures are provided in Table 2. All pairs of variables were correlated except for those involving high-frequency discrimination or Raven's test.

Figure 1 illustrates descriptive statistics for the outcome measures separately for the musically trained and untrained groups. A Multivariate Analysis of Variance (MANOVA) was initially used to determine whether the two groups differed across the set of six outcome measures. In general, musically trained participants outperformed their untrained counterparts, $F(6, 33) = 2.54$, $p < .05$. Follow-up t-tests examined whether advantages for the trained group were evident on some tasks but not on others. We used one-tailed tests because there was no reason to believe that the untrained group would be better on any measure. Because variability was significantly greater among untrained participants on three of six tasks (i.e. processing speed, low-frequency discrimination, relative pitch – 'Twinkle'; Levene's

TABLE 2 *Pairwise correlations (*N = 40*) among the outcome measures (FD = frequency discrimination, RP = relative pitch)*

	FD Low	FD High	RP Twinkle	RP Happy	Raven's Test
Processing speed	.63**	.07	.58**	.68**	.03
FD-Low		–.10	.35*	.44**	–.16
FD-High			.05	.01	–.25
RP-Twinkle				.67**	–.13
RP-Happy					–.20

*$p < .05$, **$p < .01$

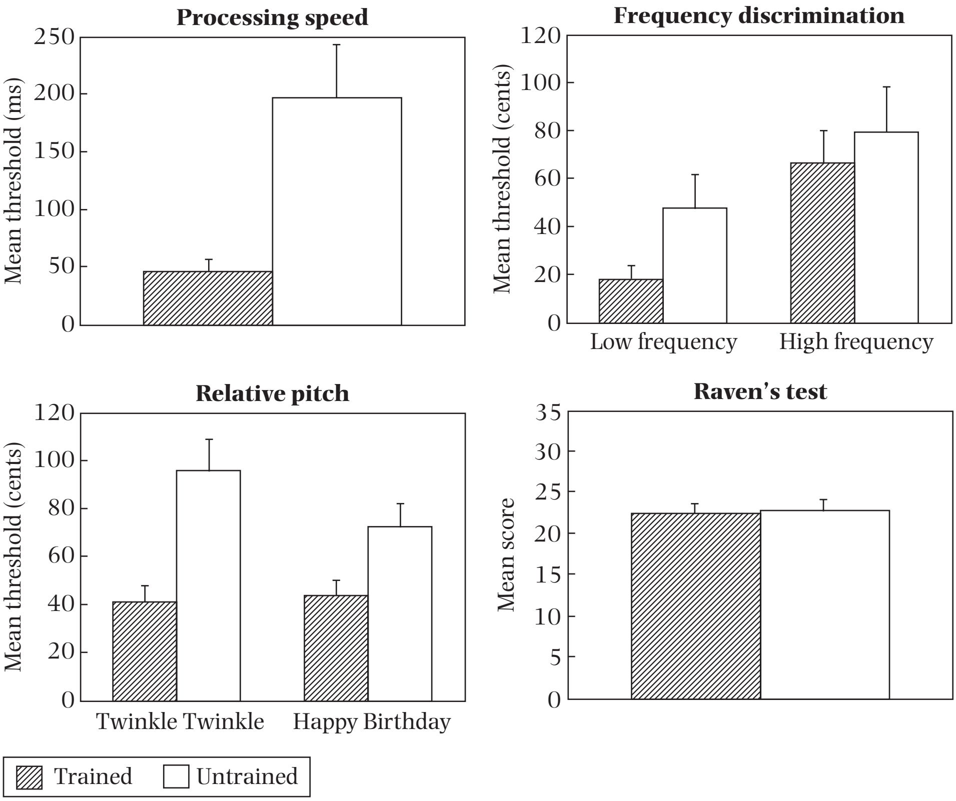

FIGURE 1 *Mean scores on the pitch-processing tasks and Raven's test for musically trained and untrained participants. Error bars are standard errors. Lower scores indicate better performance on the pitch-processing tasks. Higher scores are better on Raven's test.*

test, *p*s < .01), unequal variance *t*-tests (with non-integer degrees of freedom) were used in these instances.

On the processing-speed task, musically trained participants (*M* = 41.33 ms, *SD* = 34.44) performed well with tones of shorter duration compared to untrained participants (*M* = 181.58 ms, *SD* = 220.25), *t*(19.93) = 2.81, *p* < .01, *d* = .89. On average, trained participants performed well with tones briefer than one-twentieth of a second. Untrained participants required almost one-fifth of a second to perform equivalently.

On the low-frequency discrimination task, compared to untrained participants (*M* = 48.07 cents, *SD* = 62.22), trained participants (*M* = 18.36 cents, *SD* = 23.28) could identify the pitch contour of two tones separated by smaller differences in frequency, *t*(24.22) = 2.00, *p* < .05, *d* = .63. Trained participants performed well with tones separated by less than one-fifth of a semitone. Untrained participants required a difference of almost half a semitone. On the high-frequency discrimination task, the musically trained (*M* = 66.25 cents, *SD* = 63.23) and untrained (*M* = 79.88 cents, *SD* = 84.45) groups performed similarly and there was considerable variability within groups, *d* = .18. On average, trained and untrained groups, respectively, required a difference of two-thirds and four-fifths of a semitone to perform well.

On the relative-pitch tasks, musically trained participants ($M = 41.63$ cents, $SD = 26.00$) outperformed their untrained counterparts ($M = 95.67$ cents, $SD = 58.51$) at detecting mistunings to 'Twinkle Twinkle', $t(26.22) = 3.77$, $p < .01$, $d = 1.19$. A mistuning of approximately 40 cents was apparent to trained participants. Untrained participants required a difference of almost 1 semitone to respond with similar accuracy. Likewise, for 'Happy Birthday', trained participants ($M = 43.92$ cents, $SD = 27.34$) outperformed untrained participants ($M = 72.79$ cents, $SD = 41.18$), $t(38) = 2.61$, $p < .01$, $d = .83$. As with 'Twinkle Twinkle', the trained group required a mistuning of approximately 40 cents to perform well on the task, whereas the untrained group needed approximately 70 cents. Paired *t*-tests (two-tailed) confirmed that the untrained group required smaller mistunings for 'Happy Birthday' than for 'Twinkle Twinkle', $t(19) = 2.18$, $p < .05$, whereas the trained group performed similarly across melodies.[1]

Finally, the two groups did not differ on Raven's test, $d = .04$. In fact, the mean level of performance for the untrained group ($M = 22.80$ correct, $SD = 5.46$) was slightly higher than it was for the trained group ($M = 22.60$, $SD = 5.42$). As shown in Table 2, four of five correlations between Raven's test (higher scores = better performance) and the pitch-processing measures (lower scores = better performance) were negative but none was significant.

Discussion

We examined whether musical training predicted performance on measures of pitch processing and *g*. When the pitch-processing tasks comprised tones in a pitch range typical of music (i.e. in the two octaves above middle C), the musically trained group outperformed their untrained counterparts. Specifically, trained listeners could better identify the pitch contour of two consecutive tones of briefer duration, they could better identify the contour of two consecutive tones separated by smaller differences in pitch, and they could detect smaller mistunings to familiar melodies. When the task involved much higher tones, the advantage for musically trained participants disappeared. In fact, differences between trained and untrained groups on the high-frequency discrimination task accounted for less than 1 percent of the variance in the data. The groups also performed virtually identically on our measure of *g*, with differences between groups accounting for less than 1 percent of the variance.

The results suggest that effects of music lessons on pitch processing are broad when the tasks comprise tones in the typical pitch range for music. A null finding in the case of the high-frequency discrimination task is impossible to interpret unequivocally, however, because the sample size may have been too small or the task may have been insensitive to group differences. The latter possibility is unlikely, however, because the identical task revealed group differences with tones at a lower frequency. Although the relatively small sample size limited the power to detect a true difference between groups, if the observed effect size ($d = .18$) was an accurate reflection of the true effect size, a sample of almost 800 participants ($n = 383$ per group, one-tailed test; Faul, Erdfelder, Lang, & Buchner, 2007) would be required to reveal a significant effect with 80 percent probability or greater.

A similar analysis of statistical power helps to contextualize the null finding for Raven's test. The *untrained* group performed slightly *better* on this test. If we assume that sampling error gave us an anomalous result and that the true effect is in the opposite direction *and* thrice the magnitude, a sample of more than 1,700 undergraduates (n = 860 per group, d = .12, one-tailed test) would be required to find a significant advantage for the trained group with 80 percent probability or greater (Faul et al., 2007). In sum, although firm conclusions regarding the high-frequency discrimination task and Raven's test are precluded, if there is indeed a true advantage for musically trained undergraduates on either test, such an advantage is likely to be trivial. It is also possible that because of their relatively extensive musical training, the music group may have been quite different from typical psychology undergraduates who show a positive association between music lessons and mental abilities (Schellenberg, 2006b). Rather, the trained group may have been more like 'real' musicians, who show no general intellectual advantage over musically untrained participants (Brandler & Rammsayer, 2003; Helmbold et al., 2005). A separate issue is that the range of intellectual ability was restricted in our sample of undergraduates. An association between music training and Raven's test could be evident in a larger, more representative sample. Nonetheless, most previous findings of associations between music lessons and intellectual abilities in adulthood come from samples of undergraduates (for reviews see Schellenberg, 2005, 2006a) and it was precisely these findings that we sought to explain.

Although 'Twinkle Twinkle' and 'Happy Birthday' would have been highly familiar to both trained and untrained participants, only trained listeners' memory for the tunes included learned names for pitch *intervals* (exact distances) between consecutive tones (e.g. the second and third tones of 'Twinkle Twinkle' are separated by 7 semitones or a *perfect fifth*). Across modalities (e.g. audition, vision, olfaction), verbal encoding of stimuli leads to more detailed mental representations and enhanced memory (Bartlett, 1977; Bartlett, Till, & Levy, 1980; Bower & Holyoak, 1973; Perkins & Cook, 1990; Schab, 1991), which could help to explain the advantages for the trained group on our relative-pitch tasks. By contrast, for untrained listeners, better detection of mistunings to the tonic pitch ('Happy Birthday') than to the dominant ('Twinkle Twinkle') is likely to stem from passive exposure. Tonic pitches are used more frequently than other pitches in musical contexts (Krumhansl, 1990).

Bigand and Poulin-Charronnat (2006) proposed that the effects of musical training on music cognition are evident when listeners are asked to make *explicit* rather than *implicit* judgments about musical stimuli. In line with their view, our pitch-processing tasks required explicit judgments, and the effects of music lessons were significant whenever the stimuli comprised tones in the typical musical range. For tasks with musical relevance, music lessons appear to promote more detailed and analytic listening. The present design was a natural experiment, however, so we cannot exclude the possibility that the direction of causation was reversed: individuals with better pitch-processing abilities may be more likely to take music lessons. Such pre-existing differences could be genetic in origin. Genetics plays a role in memory for relative pitch (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001) as it does in the acquisition of absolute pitch (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Baharloo, Service, Risch, Gitschier, & Freimer, 2000).

Limitations and future directions

Our findings suggest that general associations between music lessons and non-musical cognitive functioning stem from individual differences in psychological mechanisms distinct from *g*. One candidate set of mechanisms is *executive function* (Hannon & Trainor, 2007; Schellenberg & Peretz, 2008). Executive function is predictive of performance on many IQ tests and subtests (Ardila, Pineda, & Rosselli, 2000; Carlson, Moses, & Breton, 2002; Hongwansihkul, Happaney, Lee, & Zelazo, 2005) but at least partly independent of *g*. For example, patients with frontal-lobe damage (Hebb, 1945) and children with attention-deficit/hyperactivity disorder or mild forms of autistic spectrum disorder (Pennington & Ozonoff, 1996) can have poor executive function yet perform in the normal range on tests that load highly on *g*. More importantly, unlike *g*, executive function can be improved readily through training (Zelazo, Carlson, & Kesek, 2008). Music lessons may be one type of training that leads to such improvements. Future research could attempt to draw definitive conclusions about the direction of causation between music lessons and cognitive abilities, and whether individual differences in executive function help to explain these associations. Larger and more representative samples of musically trained and untrained participants would help to ensure that interpretations of the findings would not be constrained by the limitations of the present study.

NOTES

Our research was funded by the Natural Sciences and Engineering Research Council of Canada. Will Huggon provided assistance in testing participants, data entry, and preliminary data analysis.

1. For each pitch-processing task, we examined the performance of musically trained participants in more detail by conducting comparisons between those whose primary instrument had either fixed (piano or xylophone) or variable (violin or guitar) pitch. Although the variable-pitch group had lower thresholds (i.e. better performance) on four of the five tasks (all but high-frequency discrimination), the difference between the fixed- and variable-pitch groups was not significant for any comparison, *p*s > .2. Because sample sizes and statistical power were low for these tests, significant differences may emerge with larger samples.

REFERENCES

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *15*, 31–36.

Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J., & Freimer, N.B. (1998). Absolute pitch: An approach for identification of genetic and nongenetic components. *American Journal of Human Genetics*, *62*, 224–231.

Baharloo, S., Service, S.K., Risch, N., Gitschier, J., & Freimer, N.B. (2000). Familial aggregation of absolute pitch. *American Journal of Human Genetics, 67*, 755–758.

Bartlett, J.C. (1977). Effects of immediate testing on delayed retrieval: Search and recovery operations with four types of cue. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 719–732.

Bartlett, J.C., Till, R.E., & Levy, J.C. (1980). Retrieval characteristics of complex pictures: Effects of verbal encoding. *Journal of Verbal Learning & Verbal Behavior, 19*, 430–449.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100,* 100–130.

Bors, D.A., & Stokes, T.L. (1998). Raven's Advanced Progressive Matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement, 58*, 382–398.

Bower, G.H., & Holyoak, K. (1973). Encoding and recognition memory for naturalistic sounds. *Journal of Experimental Psychology, 101*, 360–366.

Brandler, S., & Rammsayer, T.H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music, 31*, 123–138.

Carlson, S.M., Moses, L.J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*, 404–431.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271.

Deary, I.J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology, 31*, 237–250.

Deary, I.J., Spinath, F.M., & Bates, T.C. (2006). Genetics of intelligence. *European Journal of Human Genetics, 14,* 690–700.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*, 1969–1972.

Faul, F., Erdfelder, E., Lang, A-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Hamel, R., & Schmittmann, V.D. (2006). The 20-minute version as a predictor of the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 66*, 1039–1046.

Hannon, E.E., & Trainor, L.J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Science, 11*, 466–472.

Hebb, D.O. (1945). Man's frontal lobes. *Archives of Neurology & Psychiatry, 54*, 10–24.

Helmbold, N., Rammsayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences, 26*, 74–85.

Hongwanishkul, D., Happaney, K.R., Lee, W.S.C., & Zelazo, P.D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology, 28*, 617–644.

Howe, M.J.A., Davidson, J.W., & Sloboda, J.A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences, 21*, 399–442.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology, 44*, 476–490.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioural and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*, 199–211.

Marques, C., Moreno, S., Castro, S.L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*, 1453–1463.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*, 712–723.

Orsmond, G.I., & Miller, L.K. (1999). Cognitive, musical and environmental correlates of early music instruction. *Psychology of Music, 27*, 18–37.

Pennington, B.F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51–87.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688–691.

Perkins, J., & Cook, N.M. (1990). Recognition and recall of odours: The effects of suppressing visual and verbal encoding processes. *British Journal of Psychology, 81*, 221–226.

Petrill, S.A., Lipton, P.A., Hewitt, J.K., Plomin, R., Cherny, S.S., Corley, R., & DeFries, J.C. (2004). Genetic and environmental contributions to general cognitive ability through the first 16 years of life. *Developmental Psychology, 40*, 805–812.

Plomin, R., & Spinath, F.M. (2004). Intelligence: Genetics, genes, and genomics. *Journal of Personality and Social Psychology, 86*, 112–129.

Raven, J., Raven, J.C., & Court, J.H. (1998). *Advanced progressive matrices*. San Antonio, TX: Harcourt Assessment.

Schab, F.R. (1991). Odor memory: Taking stock. *Psychological Bulletin, 109*, 242–251.

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

Schellenberg, E.G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science, 14*, 317–320.

Schellenberg, E.G. (2006a). Exposure to music: The truth about the consequences. In G.E. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 111–134). Oxford: Oxford University Press.

Schellenberg, E.G. (2006b). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology, 98*, 457–468.

Schellenberg, E.G., & Peretz, I. (2008). Music, language, and cognition: Unresolved issues. *Trends in Cognitive Sciences, 12*, 45–46.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341–349.

Stalinski, S.M., Schellenberg, E.G., & Trehub, S.E. (2008). Developmental changes in the perception of pitch contour: Distinguishing up from down. *Journal of the Acoustical Society of America, 124*, 1759–1763.

Thompson, W.F., Schellenberg, E.G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*, 46–64.

Trainor, L.J. (2005). Are there critical periods for musical development? *Developmental Psychobiology, 46*, 262–278.

Trimmer, C.G., & Cuddy, L.L. (2008). Emotional intelligence, not music training, predicts recognition of speech prosody. *Emotion*, *8*, 838–849.

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

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – Third Edition*. San Antonio, TX: Psychological Corporation.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.

Zelazo, P.D., Carlson, S.M., & Kesek, A. (2008). The development of executive function in childhood. In C.A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed., pp. 553–574). Cambridge, MA: MIT Press.

E. GLENN SCHELLENBERG is a Professor of Psychology at the University of Toronto Mississauga. One of his research areas focuses on cognitive development and its relation to music. He is Secretary of the Society for Music Perception and Cognition (2007–2009).
Address: Department of Psychology, University of Toronto Mississauga, Mississauga, Ontario, Canada L5L 1C6. [email: g.schellenberg@utoronto.ca]

SYLVAIN MORENO received his PhD in Neuroscience in 2006 from the Institut de Neurosciences Cognitives de la Méditerranée (INCM) in Marseille, France. He is currently a Post Doctoral Fellow at the Rotman Research Institute Baycrest and the Cognitive Development Laboratory at York University in Toronto, Canada. His research focuses on brain plasticity and cognitive development.
Address: The Rotman Research Institute Baycrest, 3560 Bathurst Street, 914 VS, Toronto, Ontario, Canada M6A 2E1. [email: smoreno@rotman-baycrest.on.ca]