IS THERE AN ASIAN ADVANTAGE FOR PITCH MEMORY?

E. GLENN SCHELLENBERG AND SANDRA E. TREHUB
University of Toronto, Mississauga, Ontario, Canada

ABSOLUTE PITCH (AP) IS THE ABILITY TO IDENTIFY OR produce a musical note in isolation. As traditionally defined, AP requires accurate pitch memory as well as knowledge of note names. The incidence of AP is higher in Asia than it is in North America. We used a task with no naming requirements to examine pitch memory among Canadian 9- to 12-year-olds of Asian (Chinese) or non-Asian (European) heritage. On each trial, children heard two versions of a 5-s excerpt from a familiar recording, one of which was shifted upward or downward in pitch. They were asked to identify the excerpt at the original pitch. The groups performed comparably, and knowledge of a tone language did not affect performance. Nonetheless, Asians performed better on a test of academic achievement. These results provide no support for the contribution of genetics or tone-language use to cross-cultural differences in pitch memory.

Received April 15, 2007, accepted October 26, 2007.

Key words: pitch memory, absolute pitch, cross-cultural differences, music and children, academic achievement

ABSOLUTE PITCH (AP) IS THE ABILITY TO IDENTIFY or produce isolated musical tones (e.g., middle C, concert A) accurately, rapidly, and effortlessly (for reviews see Deutsch, 2002, 2006; Takeuchi & Hulse, 1993; Ward, 1999; Zatorre, 2003). It requires precise memory for the pitch of particular tones (pitch *memory*) as well as knowledge of conventional note names (pitch *labeling*; Levitin, 1994; Zatorre, 2003). AP is extremely rare, its estimated incidence ranging from 1 in 1,500 to 1 in 10,000 (Bachem, 1955; Profita & Bidder, 1988; Takeuchi & Hulse, 1993). It is often regarded as a “musical gift” (Ward, 1999) despite evidence of its interference with *relative pitch* (RP; Miyazaki, 1992, 1993, 2004; Miyazaki & Rakowski, 2002), which is central to music identification, reproduction, and memory.

The rarity of AP stems, in part, from its restriction to those with music training. AP requires knowledge of names for notes, which is lacking in untrained individuals. Despite such knowledge, most musicians do *not* possess AP, although the odds of AP increase dramatically if music training begins by 6 or 7 years of age (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Brown, Sachs, Cammuso, & Folstein, 2002; Deutsch, Henthorn, Marvin, & Xu, 2006; Gregersen, Kowalsky, Kohn, & Marvin, 1999; Miyazaki, 1988; Sergeant, 1969). Music lessons that begin later rarely lead to AP. Because early music training is the best predictor of AP, researchers have speculated about a critical period for its acquisition (e.g., Chin, 2003; Deutsch, 2002, 2006; Levitin & Zatorre, 2003; Takeuchi & Hulse, 1993; Trainor, 2005; Vitouch, 2003). Some evidence is consistent with this idea. For example, young children are better than adolescents (Crozier, 1997) and adults (Russo, Windell, & Cuddy, 2003) at learning to recognize, identify, and produce a single “special note” (e.g., concert A).

One school of thought holds that AP processing predominates in early childhood, with RP processing becoming dominant in middle childhood because of its greater efficacy for day-to-day music processing (e.g., Deutsch, 2002; Saffran, 2003; Saffran & Griepentrog, 2001; Takeuchi & Hulse, 1993). Those with early, intensive music training during the putative critical period are thought to maintain their AP skills even as they acquire RP skills. One problem with the early-exposure account, however, is that relatively few individuals who start music lessons by age 6 have AP as adults.

If early music lessons do not guarantee AP, what other factors might be implicated? Genetics could play a role, as it does with RP. In a large twin study of RP (almost 300 twin pairs), participants were asked to identify correct and incorrect renditions of familiar tunes (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001). Performance was more highly correlated for identical than for fraternal twins, resulting in a heritability estimate of approximately .75. Although there are no twin studies of AP, two lines of evidence are consistent with genetic contributions (Zatorre, 2003). The first involves a greater incidence of AP in first-degree relatives. Even when music training is held constant so that all participants began music lessons by

Music Perception VOLUME 25, ISSUE 3, PP. 241–252, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2008.25.3.241

6 years of age, the likelihood of AP is 8 to 15 times greater if a sibling has AP than otherwise (Baharloo, Service, Risch, Gitschier, & Freimer, 2000; see also Gregersen et al., 1999; Gregersen, Kowalsky, Kohn, & Marvin, 2000). Because siblings share a similar environment aside from their 50% genetic overlap, this evidence is suggestive but far from conclusive (Ward, 1999). Nonetheless, AP has a bimodal distribution (Athos et al., 2007), which Drayna (2007, p. 14549) attributes to the influence of "single major gene effects" operating in Mendelian (i.e., dominant-recessive) fashion.

The second line of evidence for genetic contributions comes from cross-cultural differences in the incidence of AP, specifically, higher incidence among Asians than among non-Asians. In one sample of over 2500 music students at conservatories or colleges in the United States, 32% of those with Asian ancestry had AP (by self-report) compared to 7% of non-Asians (Gregersen et al., 1999). In another sample of over 1000 music students in the United States, AP was reported by almost half (48%) of the students with Asian ancestry, but by only 9% of those with European ancestry (Gregersen et al., 2000). Elevated incidence was evident in each of the major Asian subgroups (Chinese, Japanese, Korean), with reports of AP ranging from one-quarter (Japanese) to almost two-thirds (Chinese) of music students. In a study involving objective AP measures (rather than self-reports) of students from music conservatories in China and the United States, the incidence of AP was approximately three times greater in the Chinese sample (Deutsch et al., 2006).

"Genetic differences may offer one possible explanation" (Zatorre, 2003, p. 694), perhaps because Asians "have a higher prevalence of AP susceptibility genes" (Gregersen et al., 1999, p. 912). Another possibility implicates cultural differences (e.g., Henthorn & Deutsch, 2007; Trehub, Schellenberg, & Nakata, in press), including type of music lessons, age of onset of lessons, and instructional focus on AP. Even when some of these factors are taken into account, the incidence of AP remains elevated among Asians (Deutsch et al., 2006; Gregersen et al., 2000). In short, differences in the incidence of AP in Asian and non-Asian samples are substantial but poorly understood, as is the etiology of AP more generally (Deutsch, 2002, 2006; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993; Ward, 1999; Zatorre, 2003). The Asian advantage could be genetic or cultural in origin (or both), reflecting differences in pitch memory or pitch labeling (or both). According to Zatorre (2003), a fruitful approach in this regard would be to study pitch memory in isolation from pitch labeling, an approach that we adopt here.

Tone-language exposure has also been offered as an explanation of the elevated incidence of AP among individuals of Asian ancestry (Deutsch, 2002, 2006; Deutsch, Henthorn, & Dolson, 2004). Across the world, there are more speakers of tone languages (e.g., Cantonese, Mandarin, Thai, Vietnamese) than of non-tone languages (e.g., English, French, German, Russian). In tone languages, the same word (e.g., /ma/) can have up to five different meanings depending on the tones used (e.g., high, low, rising). For children acquiring tone languages, intensive exposure to a limited set of pitches and their names (neither corresponding to musical pitches) occurs very early in life, in line with the proposed critical period for AP (Deutsch, 2002, 2006; Deutsch et al. 2004). As with early music lessons, however, most speakers of tone languages do not have AP. Nonetheless, individual speakers tend to produce the tones of their language at nearly identical pitch levels when they read lists of words on different occasions (Deutsch et al., 2004). Such consistency could stem, at least in part, from motor memory or limited vocal range. Even if lexical tones are relatively stable within individuals, they vary from one speaker to another, with larger differences between male and female speakers. Although acquisition of a tone language may promote attention to pitch level, it is important to note that decoding lexical tones involves RP processing, with correct classification of individual tones depending on the pitches of surrounding tones as well as the speaker's characteristic pitch range. In other words, the perceptual template for lexical tones must be flexible, much like the template for recognizing the same tune at different pitch levels. Just as AP possessors engage in RP processing, tone-language speakers may use some form of *absolute* processing for producing tones and *relative* processing for decoding the tones of other speakers.

The prevailing wisdom is that non-AP possessors have relatively poor long-term memory for the pitch level of music (e.g., Burns, 1999; Krumhansl, 2000). Nevertheless, it is important to distinguish pitch memory, which could be relatively widespread, from knowledge of pitch labels, which is restricted to individuals with music training (Zatorre, 2003). When experimental tasks are ecologically valid and exclude note labeling, untrained listeners demonstrate surprisingly good pitch memory. For example, they sing the same familiar songs at similar pitch levels from one occasion to another (Halpern, 1989), varying by less than 1 semitone in many cases (Bergeson & Trehub, 2002). When they sing as if accompanying a familiar recording, their performances are often within two semitones of the original vocal portion (Levitin, 1994).

Perceptual tasks that exclude labeling provide evidence that consistency in production does not stem solely from motor memory or restricted vocal range (Halpern, 1989). For example, tasks using tones with ambiguous pitch height but unambiguous chroma ("Shepard tones") reveal individual differences in perceptual illusions involving the chroma circle (C, C#, D, D#, and so on). Within individuals, one half of the circle is typically perceived as higher in pitch than the other half, even though there is no physical difference in pitch height (i.e., for a review see Deutsch, 1998). These findings implicate implicit but inexact memory for pitch chroma. Evidence of more accurate pitch memory comes from listeners with music training without AP, who can identify whether familiar pieces from the classical repertoire are in the notated key or transposed upward or downward by one semitone (Terhardt & Seewann, 1983; Terhardt & Ward, 1982).

Schellenberg and Trehub (2003) devised a similar task for untrained listeners. On each trial of their task, a familiar recording was presented twice: once at the original pitch level and once shifted upward or downward in pitch. Listeners judged which of the two excerpts (first or second) was presented at the original pitch level. Adults performed above chance levels (approximately 70% correct, chance = 50%) when the altered excerpt was shifted by two semitones. Performance remained above chance (approximately 60%) for pitch shifts of one semitone, which is the smallest meaningful pitch difference in Western music. Naturally, the researchers needed to rule out the possibility that audible artifacts of digital pitch shifting accounted for listeners' success in distinguishing the original from the pitch-shifted recordings. In a control experiment (Schellenberg & Trehub, 2003, Experiment 2), they pitch shifted *unfamiliar* recordings by two semitones and asked participants to identify the original. The unfamiliar recordings were closely matched to the familiar recordings from the original experiment in terms of genre, timbre, composer, and recording quality. Participants performed at chance levels (49% correct) on the unfamiliar recordings, confirming the absence of extraneous cues to pitch shifting. Ilie (2006) conducted a direct test of the efficacy of the pitch-shifting software by requiring adults to rate excerpts from instrumental recordings on the naturalness or artificiality of their sound quality. Ratings were unrelated to the presence or magnitude (one, two, or three semitones) of pitch shifting.

Does the Asian advantage in AP extend to Schellenberg and Trehub's (2003) test of pitch memory, which has no labeling requirements? That seems to be the case, although the source of the advantage remains unclear. Trehub et al. (in press) found that Japanese 4- to 7-year-olds outperformed same-age Canadian children on memory for the original pitch level of theme songs from children's television programs. It is likely that the Japanese children had less exposure to the target music than did their Canadian counterparts because the theme songs from children's TV shows are replaced much more frequently in Japan than in North America. As a result, this test of a Japanese advantage was especially conservative. Incidentally, the Canadian children performed comparably to the Canadian adults tested by Schellenberg and Trehub (2003). Even 7-month-olds exhibit memory for the pitch level of lullabies (unaccompanied vocal renditions) after 6 min of daily exposure for two weeks (Volkova, Trehub, & Schellenberg, 2006). In short, memory for pitch level, in the absence or labeling, seems to be relatively stable across development.

In the present investigation, we examined the possibility of genetic contributions to the Asian advantage for pitch memory. In previous research that reported such an advantage (Trehub et al., in press), genetics and culture were completely confounded because the Asian children were growing up in Japan whereas the non-Asian children were growing up in Canada. In the present study, the participants were 9- to 12-year-olds who were growing up in Canada. Half of the children were of Asian (Chinese) ancestry; the others were of non-Asian (European) ancestry. All of the children spoke English fluently and attended English-language schools. Recruiting children of different ancestry from a single region (greater Toronto area) reduced the usual cultural differences between groups while preserving the genetic differences. Presumably, the Asian children were influenced by their parents' culture of origin, but such influences would be reduced considerably compared to children reared in Asia. Many of the Asian children were bilingual, speaking Mandarin or Cantonese (tone languages) at home, which allowed us to examine the possibility of links between tone-language experience and pitch memory. Because the incidence of AP appears to be particularly high among Chinese musicians (Deutsch et al., 2006; Gregersen et al., 2000), recruiting children from Chinese families was assumed to maximize the likelihood of detecting differences in pitch memory between the Asian and non-Asian samples.

Because none of the children had AP, we assumed that pitch labeling would be equivalent and essentially nonexistent in both groups. We predicted, however, that both groups of children would recognize the pitch level of familiar music at above-chance levels based on previous evidence from adults (Schellenberg & Trehub, 2003), children (Trehub et al., in press), and infants (Volkova et al., 2006). The available literature precluded clear predictions about between-group comparisons. Nonetheless, each of the three possible outcomes would improve our knowledge of the development and etiology of AP. A general Asian advantage on our task (i.e., independent of tone-language exposure) would be consistent with a genetic origin for elevated rates of AP among Asians (Gregersen et al., 2000; Zatorre, 2003), and with genetic contributions to pitch memory. A performance advantage restricted to tone-language speakers would implicate linguistic experience (Deutsch, 2002, 2006; Deutsch et al., 2004, 2006). Finally, similar performance by Asian and non-Asian children would be consistent with genetic and/or cultural (but non-linguistic) factors contributing to the pitch-labeling component of AP. The same null effect could also stem from some kind of interaction between genes and the environment that becomes evident only with appropriate experience during a critical time frame.

The third possible outcome involves a failure to reject the null hypothesis, which is problematic. Failure to find differences between groups could reflect lack of statistical power or sampling problems rather than a genuine null effect. Mindful of this possibility, we included a nonmusical task for which differences between our two groups would be expected. Children of Asian origin typically outperform those of non-Asian origin on tests of academic achievement (e.g., Stevenson & Stigler, 1992; Stevenson, Stigler, Lee, Lucker, Kitamura, & Hsu, 1985), especially mathematics (e.g., Geary, 1996; Geary, Salthouse, Chen, & Fan, 1996; Stevenson, Chen, & Lee, 1993), even when both groups are raised in the same country (Chen & Stevenson, 1995). Accordingly, we administered a test of academic achievement with the goal of replicating the established findings. Successful replication would indicate that sample selection and size were sufficient for the detection of group differences.

A final goal was to explore the possibility that multiple trials with the same musical excerpts, which entail repeated exposure to the same foils (i.e., the incorrect or pitch-shifted excerpts in this case), could interfere with listeners' memory for the original pitch. As Tulving (1987) notes, information in short-term memory can inhibit the use of related information in long-term memory. Accordingly, we expected that performance would deteriorate from the first trial with each excerpt to subsequent trials, as it did for the adults tested previously on a similar task (Schellenberg & Trehub, 2003).

Method

Participants

The participants were 70 English speaking Canadian children 9 to 12 years of age ($M = 10$ years, 9 months, $SD = 10$ months), none of whom had AP. There were 35 children of Asian (Chinese) descent, 26 of whom still used their first language, either Cantonese or Mandarin, at home. In other words, these 26 children were fluent, native speakers of a tone language. The 35 non-Asian participants were of European ancestry. The Asian and non-Asian children were recruited so that the groups were matched on sex, age, and duration of music lessons. Both groups had 18 boys and 17 girls, and both groups had the same mean age and standard deviation as the combined sample.

The children had taken private music lessons for an average of 17 months ($SD = 19$ months), but the distribution was positively skewed, with a median of 12 months, and 20 children had no lessons. On average, the Asian and non-Asian children had 19.5 and 14.0 months of music lessons, respectively (SDs = 15.8 and 21.6, respectively), $p > .2$. Those with some lessons began their training, on average, at 7 years, 10 months of age ($SD = 25$ months). Because early onset of music lessons is the best predictor of AP (e.g., Vitouch, 2003), we divided the children into those with or without early lessons using four different cutoff points. One Asian and 2 non-Asians started music lessons before age 5, 5 children in both groups started before age 6, 8 Asians and 6 non-Asians started before age 7, and 15 Asians and 14 non-Asians started before age 8. Fisher's exact test confirmed that for each cutoff point, the two groups had a similar proportion of children, ps > .7.

Because the experimental task examined long-term memory for pitch, recruitment was limited to children who regularly watched at least six television programs from which the stimuli were drawn. Most children came from families who had participated previously in on-campus studies and expressed a willingness to return. Additional Asian children were recruited from a local community organization. Testing took place over a two-year period. During the first year, 28 children (14 Asians, 14 non-Asians) were tested. The remaining 42 children (21 Asians, 21 non-Asians) were tested in the second year. An additional 10 children (5 Asian, 5 non-Asian) were recruited and tested over the two years but excluded from the final sample because they offset the equivalence between groups in music training. Excluding these children had no effect on the results reported here. After finishing the test session, each child received a $10 gift certificate from a toy or music store.

TABLE 1. Television Programs from Which the Stimuli were Excerpted.

The Amanda Show	Fairly Odd Parents	Sabrina the Teenage Witch
Beyblade	Fear Factor	Simpsons*
Boy Meets World	Friends	7th Heaven
Digimon	Hockey Night in Canada	Teen Titans*
Dragonball	InuYasha*	That's So Raven
DragonballZ	Jackie Chan	Totally Spies*
Even Stevens	KimPossible	Will and Grace
Everybody Loves Raymond	Lizzy McGuire	YuGioh

Note: *Added in the second year of testing.

Stimuli and Apparatus

To identify the television shows that this age group watched most regularly, we conducted an informal telephone survey with 50 parents of children similar in age to those in the population of interest. The stimulus set (see Table 1) comprised 5-s instrumental excerpts from the theme songs of the most popular shows, provided the themes contained at least 5 s of instrumental music. In the first year of testing, the stimulus set contained 20 excerpts. An additional four excerpts were added in the second year. Each excerpt was pitch-shifted upward and downward by two semitones using the pitch-shifting function of *ProTools 5.0* (Digidesign, Daly City, CA) music-editing software. The original excerpts (i.e., those presented at the original pitch level) were first shifted upward and then downward in pitch by one semitone, so that they underwent the same amount of pitch shifting (two semitones in total) as the altered excerpts.

Children were tested individually in a sound-attenuating booth (Industrial Acoustics Co.) while seated in front of an iMac computer. Customized software created with REALbasic was used to present stimuli and record responses. The stimuli were presented over high quality headphones.

Procedure

Each child chose six shows from our stimulus set that they watched most regularly (i.e., at least weekly). The number of times each show was selected was calculated separately for the Asian and non-Asian groups. Viewing habits were similar between groups, $r = .55, N = 24, p < .01$.

The test of pitch memory comprised four blocks of six trials. Each block consisted of one trial for each of the six pairs of excerpts, presented in random order. On each trial, children heard two successive excerpts of the same recording, with 1-s silence between excerpts. Their task was to identify whether the first or second excerpt was presented at the original or "correct" pitch (i.e., the one heard at home). Half of the original-pitch excerpts were paired with excerpts that were pitch shifted upward (higher). The others were paired with downward (lower) shifts. Assignment of excerpts to upward or downward shifts was determined randomly. The direction of the pitch shift was always the same for a particular excerpt (i.e., either upward *or* downward, not upward *and* downward) to preclude a strategy of choosing the middle (correct) pitch over the course of the testing session. Order of presentation (original-shifted or shifted-original) was counterbalanced with direction of pitch shift. Thus, the correct response for any particular excerpt was "first" in two blocks of trials and "second" in the other two blocks. Moreover, the correct response corresponded to the higher-pitched excerpt for three TV shows but to the lower-pitched excerpt for the other three. The counterbalancing ensured that response strategies of consistently choosing the first (or second) excerpt, or the higher (or lower) excerpt, led to chance levels of performance (50% correct). To prevent children from learning whether the higher or lower excerpt was correct for a particular recording, no feedback was provided.

After the pitch-memory test, the Wechsler Individual Achievement Test–Second Edition–Abbreviated (WIAT-II-A; Wechsler, 2001) was administered to the 42 children who were tested in the second year. The WIAT-II-A is a brief, individually administered test for people 6 through 85 years of age. It provides achievement scores in three subject areas (*word reading*, *numerical operations*, and *spelling;* hereafter called *reading*, *math*, and *spelling*, respectively) as well as a composite score. Age-based standard scores ($M = 100, SD = 15$) were used to compare the two groups of children. Because the assistant who tested the children was a Mandarin-English bilingual with slightly accented English, words on the spelling subtest of the WIAT-II-A were digitally recorded by a native speaker of English (female) and presented as individual sound files over loudspeakers. Otherwise, the test was administered following the instructions in the WIAT-II-A manual.

Results

Pitch Memory

For ease of interpretation, performance on the pitch-memory task (number of correct responses) was converted to percent-correct scores. Preliminary analyses revealed that pitch memory was independent of the age of participants, $r = .12$, $N = 70$, $p > .3$, which is not surprising because of the relatively restricted age range of the sample. A more interesting finding was that pitch memory had no association with duration of music lessons for the group as a whole, $r = -.02$, $N = 70$, or for Asians, $r = .08$, $N = 35$, or non-Asians, $r = -.10$, $N = 35$, analyzed separately, *p*s > .5. Perhaps the range of music lessons was too limited for the emergence of a meaningful association. Indeed, only 20 of the 70 participants had 24 months or more of lessons.

Because early music lessons are the best predictor of AP, we also compared children who started music lessons early in life with those who either had no training or began lessons at an older age. Regardless of the cutoff used (i.e., lessons by age 5, 6, 7, or 8), there was no difference in pitch memory between children with early music lessons and those with later lessons or no lessons, *p*s > .2. Again, the relatively small number of participants with early music lessons diminished the possibility of significant associations. Indeed, for each comparison, absolute levels of performance were between 1.5% and 3.5% higher among children with early lessons. Interestingly, the largest difference (3.5%) occurred between children who started lessons before age 7 and the other children. Most children with AP start music lessons before age 7 (e.g., Vitouch, 2003).

On average, performance was 70.7% correct (SD = 11.9%) on the pitch-memory task, which significantly exceeded chance levels, $t(69) = 14.55$, $p < .001$, and was virtually identical to the performance of adults tested by Schellenberg and Trehub (2003) on two-semitone pitch shifts ($M = 69.6\%$). When the two groups of children were analyzed separately, both the Asian children ($M = 71.8\%$, $SD = 10.9\%$), $t(34) = 11.79$, $p < .001$, and the non-Asian children ($M = 69.7\%$, $SD = 12.9\%$), $t(34) = 9.02$, $p < .001$, surpassed chance, as expected.

Performance is illustrated in Figure 1 as a function of ancestry and trial block. A mixed-design analysis of variance (ANOVA) with one between-subjects factor (Asian, non-Asian) and one within-subjects factor (four blocks of trials) revealed a significant main effect of blocks, $F(3, 204) = 4.71$, $p = .003$. As predicted, there was a negative linear trend, $F(1, 204) = 11.40$, $p < .001$. Performance was best in the first block, worst in the fourth and final block, and intermediate in the middle two blocks. There was no difference between groups in pitch-memory performance, however, and no interaction between groups and blocks, *F*s < 1. To conduct an analysis of statistical power to detect differences between the Asian and non-Asian groups (Howell, 2007, pp. 220-221), we used our sample statistics to estimate the population means and standard deviation. The effect size we observed ($d = .18$) would require a combined sample of almost 1000 children (484 per group) to have an 80% chance of rejecting the null hypothesis.

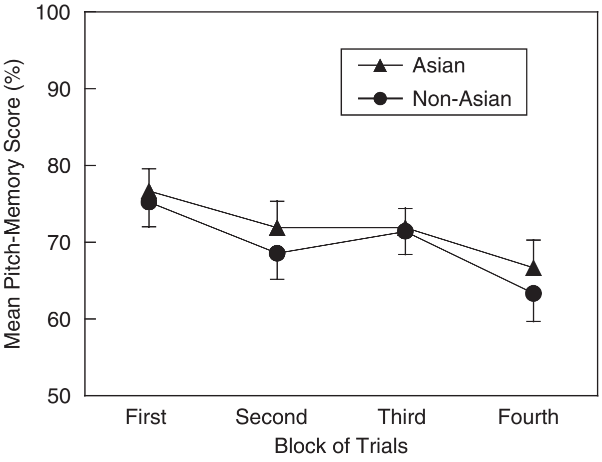

FIGURE 1. Performance of Canadian children of Asian (Chinese) and non-Asian (European) ancestry on the pitch-memory task as a function of blocks. Error bars are standard errors.

The Asian children who spoke a tone language as well as English ($M = 72.8\%$, $SD = 10.5\%$) performed no differently on the pitch-memory task than the Asian children who spoke English only ($M = 69.0\%$, SD = 12.3%), $p > .3$. When the Asian and non-Asian groups were combined, the tone-language speakers did not outperform their monolingual (English only) counterparts ($M = 69.5\%$, $SD = 12.6\%$), $p > .2$. The observed effect size in the latter case ($d = .27$) requires a sample of 430 (215 per group) to have an 80% chance of rejecting the null hypothesis. Finally, we compared the eight Asian children who spoke a tone language *and* had early music lessons (i.e., before age 7; $M = 77.1\%$, $SD = 13.0\%$) with the other 62 children ($M = 69.9\%$, $SD = 11.6\%$). Although the difference exceeded 7% and favored the Asian children, it was short of statistical significance, $p < .1$.

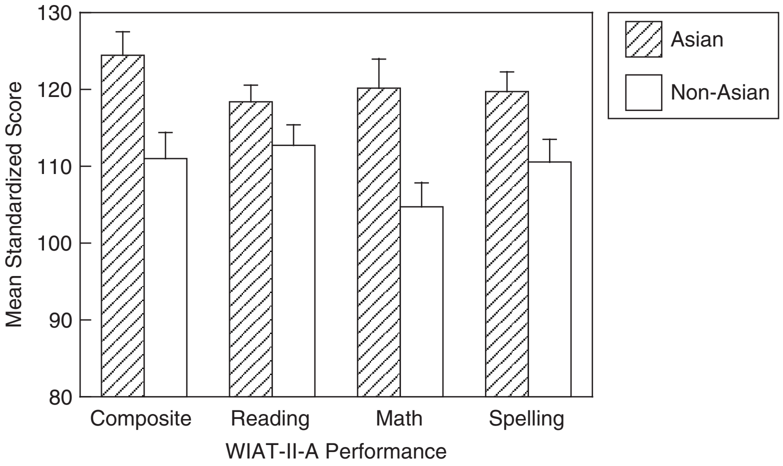

FIGURE 2. Performance of Canadian children of Asian (Chinese) and non-Asian (European) ancestry on the WIAT-II-A. Error bars are standard errors.

Academic Achievement

Descriptive statistics for scores on the WIAT-II-A are illustrated in Figure 2. For Asians, scores on the composite (overall) measure were significantly better than published (i.e., United States) norms ($M = 100$), $t(20) = 8.09$, $p < .001$, as were scores on the reading, math, and spelling subtests, $ts(20) = 8.36$, 5.32, and 7.62, respectively, $ps < .001$. For non-Asians, scores were also higher than norms on the composite measure, $t(20) = 3.25$, $p = .004$, and on the reading and spelling subtests, $ts(20) = 4.75$ and 3.58, respectively, $ps < .002$, but not on the math subtest, $p > .1$. As predicted, the Asian group outperformed the non-Asian group on the composite measure, $t(40) = 2.96$, $p = .005$, and on the math subtest, $t(40) = 3.13$, $p = .003$. Surprisingly, the Asians were also superior on the spelling subtest, $t(40) = 2.33$, $p = .025$. The groups did not differ on the reading subtest.

Discussion

We sought to further our understanding of the genesis of AP by testing whether pitch memory varies as a function of ancestry and language background. Despite reliable differences between the Asian and non-Asian groups of children in academic achievement, the groups did not differ on our test of pitch memory. Moreover, knowledge of a tone language was not associated with performance. The absence of between-group differences on the pitch-memory test cannot be attributed to floor- or ceiling-levels of performance. In general, performance was almost exactly halfway between chance (50%) and perfect (100%; see Figure 1). Moreover, our test of pitch memory is demonstrably sensitive to both within- and between-subjects manipulations predicted to affect performance. As in Schellenberg and Trehub (2003), there was a robust decrement that resulted from cumulative exposure to the pitch-shifted foils over the course of the test session. In previous research, listeners tested with two-semitone pitch shifts performed better than their counterparts tested with one-semitone shifts (Schellenberg & Trehub, 2003). When the same method was used to examine differences in pitch memory between Canadian and Japanese children in their home countries, the expected Japanese advantage was evident (Trehub et al., in press). Finally, power analyses suggested that if AP varies as a function of genetic or linguistic contributions to pitch memory, the effects are extremely small.

The present findings are consistent with Henthorn and Deutsch's (2007) reanalysis of the self-report AP data of Gregersen et al. (2000). Participants in the latter study were upper-level music students living in the United States. The original report indicated that the incidence of AP was much higher in students of Asian background than those of non-Asian background. The reanalysis demonstrated, however, that the Asian advantage disappeared when comparisons were restricted to Asians (i.e., participants self-identifying as Chinese, Japanese, or Korean) who spent their early childhood in North America. The incidence of AP for this group was no different than for non-Asians in the sample. In other words, participants who spent their early childhood in Asia accounted for the Asian advantage reported by Gregerson et al. (2000).

In sum, we found no differences in children's pitch memory that could be attributed to genetics or tone-language use. Our findings imply that dramatic differences in the incidence of AP in Asian and non-Asian populations stem either from genetic differences that affect pitch labeling, or from cultural differences (greater than those in the present study) that affect pitch memory, pitch labeling, or both components of AP. Basic associative mechanisms, such as those required to link auditory events (e.g., a tone with a fundamental frequency of 262 Hz) with arbitrary labels (e.g., middle C), could be influenced by genetics. We know that genetic factors contribute to individual differences in language acquisition and language proficiency (Stromswold, 2001). Claims of genetically based racial or ethnic differences in cognitive performance are much more contentious (Gottfredson, 2005). Nonetheless, documented Asian advantages on low-level (e.g., reaction time) and high-level (e.g., IQ) measures of intellectual function (Rushton & Jensen, 2005) allow for the possibility of genetic contributions to group differences in pitch labeling. Evidence of a normal distribution for pitch memory (Schellenberg & Trehub, 2003) and a bimodal distribution for AP (Athos et al., 2007) raises the possibility that a relatively isolated gene (Drayna, 2007) could underlie the pitch-labeling component of AP. Indeed, genetic differences between ethnic or racial groups are much more plausible for Mendelian traits (e.g., male pattern baldness, eye color), with dominant and recessive alleles, than they are for complex, polygenic traits such as intelligence (Sternberg, Grigorenko, & Kidd, 2005).

With respect to culture, what differences, other than early music lessons, could affect pitch labeling, pitch memory, or both? AP possessors identify tones faster and more accurately when the tones correspond to white-key pitches (e.g., A, B, C) than to black-key pitches (e.g., A-sharp, D-flat) and when the tones are presented on their study instrument (Athos et al. 2007; Miyazaki & Ogawa, 2006; Takeuchi & Hulse, 1991, 1993). Although these effects implicate earlier and more frequent exposure, they are likely to be comparable across cultures. Another candidate is the nature of music training (Vitouch, 2003), especially schemes for pitch naming (Levitin & Zatorre, 2003). Music training programs in Japan and various non-English speaking countries (e.g., France) use a *fixed-do* (C = *do*) system (e.g., Suzuki method), in which specific pitches retain their names regardless of their musical function (e.g., tonic, leading note) or position in a scale. Thus, scales can start on *do*, *re*, *mi*, or any other pitch. Invariant relations between specific pitches and labels could enhance the memory and labeling components of AP. By contrast, the *moveable-do* system, which is common in English-speaking countries, uses *do* to refer to the key of a musical piece and the starting pitch of a musical scale, which can assume any of 12 different pitch values (C, C#, D, D#, and so on). Although the use of fixed *do* could contribute to higher rates of AP in Japan than in North America, it cannot account for the high rates of AP in China (Deutsch et al., 2006), which uses a movable *do*. Moreover, there is no evidence to date of elevated rates of AP in European countries that use a fixed-*do* system. Even in cultures that use moveable *do*, instructors are likely to make frequent use of letter names (C, C#, D, D#, and so on) for tones during music lessons. In short, the use of fixed versus movable *do* seems unlikely to account for differences in the incidence of AP across countries.

Could differences in the incidence of AP be based on tone-language use, as proposed by Deutsch and her associates (2002, 2006; Deutsch et al., 2004, 2006)? Performance on our pitch-memory test was not significantly better for bilingual participants whose native language was a tone language. It is possible, however, that experience with specific linguistic tones (e.g., high, low, falling) that are part of a language such as Mandarin, Cantonese, Vietnamese, or Thai facilitates the acquisition of labels for specific pitches. In countries with pitch-accent languages (e.g., standard Tokyo dialect, some Korean dialects) as opposed to tone languages, the system of different linguistic pitches is much less complex, with a limited number of syllables (or morae) resulting in lexical contrasts because of their high or low pitch. In the Tokyo dialect, for example, /kaki/ refers to "oyster" when spoken with a drop in pitch (or "downstep") between the first and second syllables and "fence" when spoken with a corresponding rise in pitch. The use of pitch to signal semantic differences could enhance attention to individual pitches and pitch patterning, which, in turn, could facilitate the acquisition of pitch labels. From this perspective, one would expect a high incidence of AP among speakers of tone languages (with the most elaborate pitch-meaning mappings), moderate rates among speakers of pitch-accent languages, and the lowest rates among speakers of stress-accent languages, such as English. The available information on the incidence of AP across language groups is in line with these predictions (Gregersen et al., 2000).

Nonlinguistic cultural differences may also contribute to the acquisition of AP, as they do in academic achievement. In one study of fifth-graders (Stevenson et al., 1993), Japanese and Chinese students outperformed their American counterparts in mathematics and reading, yet American mothers were more likely to be "very satisfied" with their children's performance, and Asian mothers were more likely to be "not satisfied." The students showed similar patterns of self-ratings, with the lowest scoring American students being the most satisfied with their performance. Whereas Asian cultures stress "hard work as the route to success," American culture stresses "the importance of innate abilities" (Stevenson et al., 1993, p. 57). Other cultural differences include time in school, quality of instruction, time doing homework, and attitudes toward homework, with Asians exhibiting advantages in each case (Stevenson & Lee, 1990; Stigler, Gallimore, & Hiebert, 2000). In principle, such cultural differences could extend to TV viewing. Although we ascertained that each child participant watched each of their six selected TV shows on a weekly basis (or more regularly), cumulative exposure may have varied within and across cultures, which could have influenced performance on our test of pitch memory. One challenge for future research is to quantify exposure (e.g., TV viewing habits) or control it (e.g., supplying novel materials for specified viewing times) to ascertain the effects of factors such as hours of exposure, intensity of exposure (e.g., number of exposures per week), and recency of exposure.

Our findings add to our knowledge of memory for absolute aspects of ecologically valid stimuli. Obviously, music is primarily a relational domain because tunes are readily identifiable across transformations in pitch, tempo, and timbre. As in previous research (Bergeson & Trehub, 2002; Halpern, 1989; Levitin, 1994; Schellenberg & Trehub, 2003; Volkova et al., 2006), the findings indicate that non-AP possessors exhibit memory for pitch in the context of ecologically valid tasks and stimuli. Our results corroborate other reports of high-fidelity memory traces of meaningful music (Dowling, Tillmann, & Ayers, 2002) and speech (Goldinger, 1996). In addition to pitch, listeners' representations of music include fine-grained information about tempo (Levitin & Cook, 1996) and timbre (Schellenberg, Iverson, & McKinnon, 1999).

By replicating and extending the finding of pitch memory in untrained adults to a population of children 9 to 12 years of age, our results support the notion that the rarity and genesis of AP (Ward, 1999) are linked to pitch labeling rather than to pitch memory. Our findings also indicate that conventional means of testing musical memory—involving repeated trials with the same stimuli—interfere with the retrieval of surface features from long-term memory.

Even though knowledge of note names is essential for AP, pitch memory is relatively independent of music lessons, at least for those with very limited training. We propose that the critical period for AP acquisition stems from age-related constraints on pitch labeling that have little consequence for pitch perception or memory. The critical period for associating sensory qualities (i.e., specific pitches in this case) with verbal labels may be related to optimal periods for *fast-mapping* (e. g., Carey & Bartlett, 1978; Heibeck & Markman, 1987; Markson & Bloom, 1997) and the ability to achieve native-like competence in the sound system of a second language (e. g., Piske, MacKay, & Flege, 2001). Age-related decrements in AP acquisition are thought to have a relatively *hard* cut-off (6 or 7 years; Vitouch, 2003), but special circumstances that generate a protracted or atypical course of cognitive development such as Williams syndrome (Lenhoff, Perales, & Hickok, 2001a, 2001b) or autism (Heaton, Hermelin, & Pring, 1998; Heaton, Pring, & Hermelin, 1999; Mottron, Peretz, Belleville, & Rouleau, 1999) seem to extend the usual window of opportunity. Sensitive periods in different domains may have less to do with age than with limited expertise or neural commitment in those domains, which enables new learning to proceed without undue interference from prior learning (Johnson & Munakata, 2005; Seidenberg & Zevin, 2006).

In summary, it is likely that AP is the consequence of interactions between genetic and experiential factors that influence the ability to form arbitrary associations, such as those between note names and musical pitches. Other developmental studies of music (Hannon & Trehub, 2005a, 2005b; Lynch, Eilers, Oller, & Urbano, 1990; Schellenberg & Trehub, 1999; Trainor & Trehub, 1992, 1994; Trehub, Schellenberg, & Kamenetsky, 1999) indicate that infancy and early childhood are periods of enhanced perceptual flexibility. Young listeners with limited exposure and weak representations may acquire new and enduring associations more readily than older listeners. Development cannot be the whole story, however, or AP would result from all instances of early music training. Rather, enduring associations between musical pitches and labels seem to rely on genetic dispositions that interact with environmental factors, including appropriate exposure at the appropriate time of life. The result is a distinctive perceptual style that is reflected in unique brain structure and function (Zatorre, 2003). The present study clarifies some aspects of long-term memory for pitch level. Further examination of pitch-memory and pitch-labeling skills will improve our understanding of the determinants of AP.

Author Note

Funded by the Natural Sciences and Engineering Research Council of Canada. Jane Campbell, Joyet Chakungal, and Amelia Woo provided assistance in recruiting and testing participants, stimulus selection, and data entry.

Correspondence concerning this article should be addressed to Glenn Schellenberg, Department of Psychology, University of Toronto at Mississauga, Mississauga, ON, Canada L5L 1C6. E-MAIL: g.schellenberg@utoronto.ca

References

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences (USA), 104*, 14795-14800.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27*, 1180-1185.

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.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science, 13,* 72-75.

Brown, W. A., Sachs, H., Cammuso, K., & Folstein, S. E. (2002). Early music training and absolute pitch. *Music Perception, 19*, 595-597.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 215-264). San Diego, CA: Academic Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17-29.

Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development, 66*, 1215-1234.

Chin, C. S. (2003). The development of absolute pitch: A theory concerning the roles of musical training at an early developmental age and individual cognitive style. *Psychology of Music, 31,* 155-171.

Crozier, J. B. (1997). Absolute pitch: Practice makes perfect, the earlier the better. *Psychology of Music, 25,* 110-119.

Deutsch, D. (1998). The tritone paradox: A link between music and speech. *Current Directions in Psychological Science, 6*, 174-180.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science, 11*, 200-204.

Deutsch, D. (2006). The enigma of absolute pitch. *Acoustics Today, 2*, 11-19.

Deutsch, D., Henthorn, T., & Dolson, T. (2004). Absolute pitch, speech, and tone language: Some experiments and a proposed framework. *Music Perception, 21,* 339-356.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period (L). *Journal of the Acoustical Society of America, 119*, 719-722.

Dowling, W. J., Tillmann, B., & Ayers, D. F. (2002). Memory and the experience of hearing music. *Music Perception, 19*, 249-276.

Drayna, D. T. (2007). Absolute pitch: A special group of ears. *Proceedings of the National Academy of Sciences (USA), 104*, 14549-14550.

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

Geary, D. C. (1996). International differences in mathematical achievement: Their nature, causes, and consequences. *Current Directions in Psychological Science, 5*, 133-137.

Geary, D. C., Salthouse, T. A., Chen, G,-P., & Fan, L. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Developmental Psychology, 32*, 254-262.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166-1183.

Gottfredson, L. S. (2005). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law, 11,* 311-319.

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (1999). Absolute pitch: Prevalence, ethnic variation, and estimation of the genetic component. *American Journal of Human Genetics, 65,* 911-913.

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (2000). Early music education and predisposition to absolute pitch: Teasing apart genes and environment. *American Journal of Human Genetics, 98,* 280-282.

HALPERN, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory and Cognition, 17,* 572-581.

HANNON, E. E., & TREHUB, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, *16*, 48-55.

HANNON, E. E., & TREHUB, S. E. (2005b). Tuning into musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences (USA), 102*, 12639-12643.

HEATON, P., HERMELIN, B., & PRING, L. (1998). Autism and pitch processing: A precursor for savant musical ability? *Music Perception, 15*, 291-305.

HEATON, P., PRING, L., & HERMELIN, B. (1999). A pseudo-savant: A case of exceptional musical splinter skills. *Neurocase, 5*, 503-509.

HEIBECK, T. H., & MARKMAN, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*, 1021-1034.

HENTHORN, T., & DEUTSCH, D. (2007). Ethnicity versus early environment: Comment on 'Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment' by Peter K. Gregersen, Elena Kowalsky, Nina Kohn, and Elizabeth West Marvin [2000]. *American Journal of Medical Genetics, 143A*, 102-103.

HOWELL, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.

ILIE, G. (2006). *The role of tempo, pitch height, and intensity in the perception and experience of affect: A music and speech evaluation*. Unpublished doctoral dissertation, University of Toronto.

JOHNSON, M. H., & MUNAKATA, Y. (2005). Processes of change in brain and cognitive development. *Trends in Cognitive Sciences, 9*, 152-188.

KRUMHANSL, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159-179.

LENHOFF, H. M., PERALES, O., & HICKOK, G. (2001a). Absolute pitch in Williams syndrome. *Music Perception, 18*, 491-503.

LENHOFF, H. M., PERALES, O., & HICKOK, G. (2001b). Preservation of a normally transient critical period in a cognitively impaired population: Window of opportunity for acquiring absolute pitch in Williams syndrome. In C. A. Shaw & J. C. McEachern (Eds.), *Toward a theory of neuroplasticity* (pp. 275-287). Philadelphia: Psychology Press.

LEVITIN, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned memories. *Perception and Psychophysics, 56,* 414-423.

LEVITIN, D. J., & COOK, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception and Psychophysics, 58*, 927-935.

LEVITIN, D. J., & ROGERS, S. E. (2005). Absolute pitch: Perception, coding and controversies. *Trends in Cognitive Sciences, 9*, 26-33.

LEVITIN, D. J., & ZATORRE, R. J. (2003). On the nature of early music training and absolute pitch: A reply to Brown, Sachs, Cammuso, and Folstein. *Music Perception, 21*, 105-110.

LYNCH, M. P. EILERS, R. E., OLLER, D. K., & URBANO, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1*, 272-276.

MARKSON, L., & BLOOM, P. (1997). Evidence against a dedicated system for word learning in children. *Nature, 385*, 813-815.

MIYAZAKI, K. (1988). Musical pitch identification by absolute pitch processors. *Perception and Psychophysics, 44,* 501-512.

MIYAZAKI, K. (1992). Perception of musical intervals by absolute pitch possessors. *Music Perception, 9*, 413-426.

MIYAZAKI, K. (1993). Absolute pitch as an inability: Identification of musical intervals in a tonal context. *Music Perception, 11*, 55-72.

MIYAZAKI, K. (2004). Recognition of transposed melodies by absolute-pitch possessors. *Japanese Psychological Research, 46*, 270-282.

MIYAZAKI, K., & OGAWA, Y. (2006). Learning absolute pitch by children: A cross-sectional study. *Music Perception, 24*, 63-78.

MIYAZAKI, K., & RAKOWSKI, A. (2002). Recognition of notated melodies by possessors and non-possessors of absolute-pitch. *Perception and Psychophysics, 64*, 1337-1345.

MOTTRON, L., PERETZ, I., BELLEVILLE, S., & ROULEAU, N. (1999). Absolute pitch in autism: A case study. *Neurocase, 5*, 485-501.

PISKE, T., MACKAY, I. R. A., & FLEGE, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215.

PROFITA, J., & BIDDER, T. G. (1988). Perfect pitch. *American Journal of Medical Genetics, 29,* 763-771.

RUSHTON, J. P., & JENSEN, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235-294.

RUSSO, F. A., WINDELL, D. L., & CUDDY, L. L. (2003). Learning the "special note": Evidence for a critical period for absolute pitch acquisition. *Music Perception, 21*, 119-127.

SAFFRAN, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science, 6*, 35-47.

SAFFRAN, J. R., & GRIEPENTROG, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37,* 74-85.

SCHELLENBERG, E. G., IVERSON, P., & MCKINNON, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin and Review, 6*, 641-646.

SCHELLENBERG, E. G., & TREHUB, S. E. (1999). Culture-general and culture-specific factors in the discrimination of melodies. *Journal of Experimental Child Psychology, 74*, 107-127.

SCHELLENBERG, E. G., & TREHUB, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14*, 262-266.

Seidenberg, M. S., & Zevin, J. D. (2006). Connectionist models in developmental cognitive neuroscience: Critical periods and the paradox of success. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development: Attention and performance XXI* (pp. 315-347). Oxford, UK: Oxford University Press.

Sergeant, D. (1969). Experimental investigation of absolute pitch. *Journal of Research in Music Education, 17*, 135-143.

Sternberg, R. J., Grigorenko, E. L., & Kidd, K. K. (2005). Intelligence, race, and genetics. *American Psychologist, 60*, 46-59.

Stevenson, H. W., Chen, C., & Lee, S.-Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science, 259*, 53-58.

Stevenson, H. W., & Lee, S. (1990). Contexts of achievement. *Monographs of the Society for Research in Child Development, 55* (Serial No. 221).

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit Books.

Stevenson, H. W., Stigler, J. W., Lee, S. Y., Lucker, G. W., Kitamura, S., & Hsu, C. C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development, 56*, 718-734.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35*, 87-100.

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption and linkage studies. *Language, 77,* 647-723.

Takeuchi, A. H., & Hulse, S. H. (1991). Absolute-pitch judgments of black- and white-key pitches. *Music Perception, 9*, 27-46.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113,* 345-361.

Terhardt, E., & Seewann, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Perception, 1*, 63-83.

Terhardt, E., & Ward, W. D. (1982). Recognition of musical key: Exploratory study. *Journal of the Acoustical Society of America, 72*, 26-33.

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

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 394-402.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics, 56*, 125-132.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 965-975.

Trehub, S. E., Schellenberg, E. G., & Nakata, T. (in press). Cross-cultural perspectives on pitch memory. *Journal of Experimental Child Psychology.*

Tulving, E. (1987). Multiple memory systems and consciousness. *Human Neurobiology, 6,* 67-80.

Vitouch, O. (2003). Absolutist models of absolute pitch are absolutely misleading. *Music Perception, 21*, 111-117.

Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2006). Infants' memory for musical performances. *Developmental Science, 9*, 583-589.

Ward, W. D. (1999). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp 265-298). San Diego, CA: Academic Press.

Wechsler, D. (2001). *Wechsler individual achievement test – Second edition abbreviated*. San Antonio, TX: The Psychological Corporation.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience, 6*, 692-695.