Music Training and Nonmusical Abilities: Commentary on Stoesz, Jakobson, Kilgour, and Lewycky (2007) and Jakobson, Lewycky, Kilgour, and Stoesz (2008)

E. Glenn Schellenberg
University of Toronto

Stoesz, Jakobson, Kilgour, and Lewycky (2007) and Jakobson, Lewycky, Kilgour, and Stoesz (2008) examined whether music training is associated with nonmusical abilities. They concluded that their results provided evidence of "specific" associations between music training and local-processing abilities (Stoesz et al., 2007), and between music training and memory for verbal and visual stimuli (Jakobson et al., 2008). Closer inspection of their methods, consideration of the available literature, and a reanalysis of previous data reveal that these conclusions are debatable. Moreover, the causal direction of the observed associations could go either or both ways.

Received July 14, 2008, accepted March 18, 2009.

Key words: music training, cognitive transfer, processing style, intelligence, expertise

The possibility that music lessons have positive beneficial effects for cognitive functioning in nonmusical domains (e.g., verbal, spatial, mathematical) has become an area of intense research focus. A search of articles in *Psycinfo*[1] uncovered 13 empirical papers published in 2007 and 2008 that compared musically trained and untrained participants in terms of linguistic (Dankovicová, House, Crooks, & Jones, 2007; Marques, Moreno, Castro, & Besson, 2007; Milovanov, Tervaniemi, Takio, & Hämäläinen, 2007; Musacchia, Sams, Skoe, & Kraus, 2007) or other nonmusical (Franklin et al., 2008; Hughes & Franz, 2007; Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Lee, Lu, & Ko, 2007; Patston, Hogg, & Tippett, 2007; Patston, Kirk, Rolfe, Corballis, & Tippett, 2007; Sluming, Brooks, Howard, Downes, & Roberts, 2007; Stoesz, Jakobson, Kilgour, & Lewycky, 2007; Yong & McBride-Chang, 2007) abilities, as well as three theoretical papers (Hannon & Trainor, 2007; Patel & Iversen, 2007; Schellenberg & Peretz, 2008).

Despite this flurry of interest, central issues remain unresolved, including (1) the direction of causation and (2) whether associations between music lessons and intellectual abilities are *specific* or *general* (Schellenberg, 2005, 2006a, 2008; Schellenberg & Peretz, 2008). The former issue is considered only briefly here, with the principal focus being on the latter. Whereas the "specifists" suggest that links between music lessons and intellectual abilities are evident in some domains but not in others, the "generalist" view holds that such associations are broad, stemming from differences in general intelligence (typically measured by Full-Scale IQ: FSIQ) and extending across virtually every area of intellectual ability one chooses to study. The present commentary evaluates two recent reports (Jakobson et al., 2008; Stoesz et al., 2007) in order to determine whether claims of specific links between music training and cognitive abilities are warranted based on the evidence provided. These two reports are not alone in terms of their claims of specificity. Rather, both appeared in *Music Perception*, whose readers are likely to design future research based on the available literature.

Stoesz et al. (2007) compared undergraduates with or without music training on three nonmusic tasks: a visual search task (Group Embedded Figures Test; Witkin, Oltman, Raskin, & Karp, 1971) in Experiment 1 (hereafter Sample 1), and, in Experiment 2 (Sample 2), a visuo-spatial task (the Block Design subtest from the Wechsler Adult Intelligence Scale—Third Edition: WAIS, Wechsler, 1997) and a task that required participants to copy line drawings (Mottron, Belleville, & Menard, 1999). Jakobson et al. (2008) tested a subset of participants from Sample 2 (36 of 43; Sample 3) on measures of verbal memory (California Verbal Learning Test—Second Edition [CLVT-II], Delis, Kramer, Kaplan & Ober, 2000) and visual memory (Rey Visual Design Learning Test [RVDLT], Graves & Sarazin, 1985; Rey, 1964).

[1]A search of articles published in 2007 and 2008 was conducted in February, 2009, with keywords "music lessons" or "music training" or "musical training." Unpublished dissertations were excluded, as were articles published in languages other than English, and articles that examined differences between musically trained and untrained participants in music-related or non-cognitive abilities.

Music Perception volume 27, issue 2, pp. 139–143, issn 0730-7829, electronic issn 1533-8312 DOI:10.1525/mp.2009.27.2.139

Across samples, the trained and untrained groups were matched in terms of age and gender but musically trained participants tended to have parents with more education.

Samples 2 and 3 had an additional measure of reading ability—the North American Adult Reading Test (NAART; Blair & Spreen, 1989)—which required participants to read aloud 61 English words, each with irregular spelling (e.g., *debt*, *cellist*, *ennui*). The NAART was designed to estimate premorbid intellectual abilities in patients with dementia. It has good reliability (Spreen & Strauss, 1998; Uttl, 2002), and its validity is demonstrated by associations with the Vocabulary subtest from the WAIS in large non-clinical samples (*N*s > 300; Uttl, 2002, $r = .75$; Wiens, Bryan, & Crossen, 1993, $r = .54$). Unlike Vocabulary, the NAART can be administered by virtually anyone; each word is scored *correct* or *incorrect* based on how it is pronounced. Because performance on the NAART is correlated with FSIQ ($.4 \leq r \leq .8$; Spreen & Strauss, 1998), as are almost all measures of intellectual ability (Carroll, 1993), NAART scores can be used to estimate FSIQ by means of the regression equation. This is true for any pair of correlated variables, however, and it makes no difference in the statistical analyses which variable is used because the equation is a linear transformation of one variable (pronunciation errors on the NAART) into another (FSIQ).

In Sample 1, musically trained participants outperformed their untrained counterparts on the visual-search task. Trained participants also performed better on six measures of verbal and visual memory in Sample 3 (CLVT-II: cued recall short delay, free recall long delay, cued recall long delay; RVDLT: immediate free recall on two of five trials, delayed free recall, delayed recognition) and on the NAART in Samples 2 and 3. More importantly, the music group was superior on Block Design and at copying "impossible" figures after parental education and NAART scores were held constant (Sample 2), and on two measures of memory (CLVT-II: free recall long delay; RVDLT: delayed free recall) after NAART scores (but not parental education) were held constant (Sample 3). The authors interpreted their results as evidence of specific links between music training and local-processing abilities (Stoesz et al., 2007), and between music training and memory (Jakobson et al., 2008).

The available literature indicates, however, that music training is associated with performance on many cognitive tasks (for reviews see Schellenberg, 2005, 2006a). Indeed, Stoesz et al. (2007) noted that previous findings from their laboratory revealed advantages for participants with music training on "a number of nonmusic, perceptual, and cognitive abilities including auditory temporal processing, verbal and visual memory, and certain executive functions" (p. 160). General associations between music training and intellectual functioning were also evident in an experiment (Schellenberg, 2004) in which 6-year-olds were administered the entire Wechsler Intelligence Scale for Children—Third Edition (WISC, Wechsler, 1991), once before entering first grade and again before second grade. Children assigned randomly to music lessons in the interim had larger increases in FSIQ than children assigned to drama lessons or no lessons. These advantages extended across the 12 WISC subtests (see Schellenberg, 2005/2006, for more detailed analyses). In a follow-up correlational study (Schellenberg, 2006b), duration of formal involvement with music was associated positively with FSIQ and with many of the WISC and WAIS subtests for children and adults, respectively. For each subtest, the association with music training disappeared when individual differences in general intelligence were equated (i.e., held constant) statistically.

Considered jointly, Stoesz et al. (2007) and Jakobson et al. (2008) reported positive associations between music training and measures of (1) verbal ability (NAART), (2) visuo-spatial ability (Block Design), (3) processing style (embedded figures), (4) visuo-motor ability (copying impossible figures), (5) verbal memory (CLVT-II), and (6) visual memory (RVDLT). The authors maintained, however, that "training is associated *specifically* with enhanced local processing skills beyond any benefits it may have on verbal (or general) intelligence" (Stoesz et al., 2007, p. 162, emphasis added, parentheses in original) and that "formal music training is associated with superior performance in multiple domains of memory functioning, above and beyond any effects it may have on general intelligence" (Jakobson et al., 2008, p. 50). The crucial point is that the authors did *not* administer a test of general intelligence, only a single measure of pronunciation. Although they cited a correlation between NAART scores and FSIQ of .8, this statistic came from a clinical report of 64 patients, which stated explicitly that the association needed "to be validated on a larger and more heterogeneous sample" (Griffin, Mindt, Rankin, Ritchie, & Scott, 2002, p. 505). A review of multiple studies suggests that the correlation is somewhere between .4 and .8 (Spreen & Strauss, 1998). In a large, non-clinical sample, it was .46 (Wiens et al., 1993). Although it's clear that as FSIQ increases, so does the likelihood of pronouncing words like *caveat* and *epitome* correctly, it's equally clear that a test of pronunciation is *not* a measure of general intelligence.

Moreover, because the NAART comes from a different domain (linguistic) than the other tests administered by Stoesz et al. (2007) and Jakobson et al. (2008), NAART scores and the outcome measures had only a modicum of variance in common. In other words, much unexplained (but not "specific") variance remained in the outcome measures after individual differences in NAART scores were held constant, which maximized the likelihood that associations with music training would be evident. If general intelligence were held constant instead, there would be less unexplained variance in the outcome measures. Consequently, tests of differences between musically trained and untrained participants would be more conservative. To illustrate, individual differences in FSIQ explain 46% of the variance in Block Design, whereas performance on the Vocabulary subtest accounts for only 17% (Wechsler, 1997). The overlap between NAART scores and Block Design is even smaller (4% in Wiens et al., 1993).

A reanalysis of data from Schellenberg's (2006b, Experiment 2, $N = 150$) correlational study of undergraduates demonstrates how it is one thing to hold constant general intelligence, and another to hold constant performance on a specific subtest. Participants with or without private music lessons (following the quasi-experimental design of Jakobson et al., 2008 and Stoesz et al., 2007) were compared on Block Design while holding constant *either* Vocabulary or general intelligence (derived without Block Design).[2] When Vocabulary was held constant, the music group outperformed the nonmusic group, $F(1, 147) = 7.17, p < .01$. When general intelligence was held constant, the group effect was no longer significant, $F(1, 147) = 3.54, p > .05$. There is no reason to believe that response patterns for the copying or free-recall tasks would differ in this regard. Performance on Block Design is associated positively with copying impossible figures (Stoesz et al., 2007) and with the Working Memory index from the WAIS (Wechsler, 1997).

On the one hand, the reanalysis of Schellenberg's (2006) data replicate perfectly the results from Stoesz et al. (2007) involving music training, Block Design, and a measure of vocabulary. On the other hand, the reanalysis does not reveal evidence of a specific association between Block Design and music training. Whereas Stoesz et al. proposed a specific link between music training and local processing, Jakobson et al. (2008) proposed a specific link between music training and memory based on the finding that some of the observed simple associations (two of six) were still significant after holding constant NAART scores. Other researchers have proposed specific links between music lessons and language, mathematical, or spatial abilities (for reviews see Schellenberg, 2005, 2006a). Multiple so-called "specific" links belie each claim of specificity. In sum, the relevant behavioral literature considered as a whole (Schellenberg, 2005, 2006a), individual studies with multiple measures from different domains (Schellenberg 2004, 2006b), and the data from Stoesz et al. and Jakobson et al. all point to a very general link between music training and intellectual functioning. To date, however, this link remains poorly understood.

Although the underlying mechanisms of the general association remain elusive (Hannon & Trainor, 2007; Patel & Iversen, 2007; Schellenberg & Peretz, 2008), almost all of the available data can be explained simply: Individuals with higher IQs are more likely to take music lessons than their lower-IQ peers, and they are also likely to perform better on virtually any test of cognitive abilities. According to the rules of science, this parsimonious interpretation should be the norm rather than the exception. Yet many colleagues, including most of the authors of empirical papers published in 2007 and 2008 (cited in the first paragraph), favor the opposite interpretation, that "music training produces a range of beneficial effects on perceptual and cognitive functioning" (Stoesz et al., 2007, p. 160) and that "formal music training 'sculpts' the brain" (Jakobson et al., 2008). In their defense, they often cite the single study that allows for inferences of causation (Schellenberg, 2004), as well as studies of brain structure and function that reveal differences between participants with or without music training (for a review see Patel, 2008). Nonetheless, a single finding of small effects of music lessons on cognition should not be treated as gospel or preclude the possibility that the direction of causality also goes in the reverse direction. Brain data from quasi-experiments are even less compelling. If preexisting intellectual differences influence the likelihood of taking music lessons, such differences must be instantiated somewhere in the brain.

Author Note

Supported by the Natural Sciences and Engineering Research Council of Canada.

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

[2] FSIQ is an aggregate measure formed from Block Design and other WAIS subtests (Wechsler, 1997). In this instance, general intelligence was calculated as the score on the principal component derived from all WAIS subtests *other* than Block Design. This alternative measure was correlated almost perfectly with FSIQ, $r(148) = .97, p < .0001$.

References

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 3*, 129-136.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Dankovicová, J., House, J., Crooks, A., & Jones, K. (2007). The relationship between musical skills, music training, and intonation analysis skills. *Language and Speech, 50*, 177-225.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test—Second Edition (CVLT-II).* San Antonio, TX: Psychological Corporation.

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

Graves, R. E., & Sarazin, F. (1985). *RVDLT*. Unpublished manuscript, University of Victoria.

Griffin, S. L., Mindt, M. R., Rankin, E. J., Ritchie, A. J., & Scott, J. G. (2002). Estimating premorbid intelligence: Comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology, 17*, 497-507.

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

Hughes, C. M. L., & Franz, E. A. (2007). Experience-dependent effects in unimanual and bimanual reaction time tasks in musicians. *Journal of Motor Behavior, 39*, 3-8.

Jakobson, L. S., Lewycky, S. T., Kilgour, A. R., & Stoesz, B. M. (2008). Memory for verbal and visual material in highly trained musicians. *Music Perception*, *26*, 41-55.

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction, 17*, 336-344.

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.

Milovanov, R., Tervaniemi, M., Takio, F., & Hämäläinen, H. (2007). Modification of dichotic listening (DL) performance by musico-linguistic abilities and age. *Brain Research, 1156*, 168-173.

Mottron, L., Belleville, S., & Menard, E. (1999). Local bias in autistic subjects as evidenced by graphic tasks: Perceptual hierarchization or working memory deficit? *Journal of Child Psychology and Psychiatry, 40*, 743-755.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences (USA), 104*, 15894-15898.

Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences, 11*, 369-372.

Patston, L. L. M., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality, 12*, 262-272.

Patston, L. L. M., Kirk, I. J., Rolfe, M. H. S., Corballis, M. C., & Tippett, L. J. (2007). The unusual symmetry of musicians: Musicians have equilateral interhemispheric transfer for visual information. *Neuropsychologia, 45*, 2059-2065.

Rey, A. (1964). *L'examen clinique en psychologie* [Clinical examination in psychology]. Paris: Presses Universitaire de France.

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*, 322-325.

Schellenberg, E. G. (2005/2006). Music lessons enhance IQ: A reply to Black (2005) and Steele (2005). *Scientific Review of Mental Health Practice, 4,* 10-13.

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, UK: 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. (2008). Commentary on "Effects of early musical experience on auditory sequence memory" by Adam Tierney, Tonya Bergeson, and David Pisoni. *Empirical Musicology Review, 3,* 205-207.

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

Sluming, V., Brooks, J., Howard, M., Downes, J. J., & Roberts, N. (2007). Broca's area supports enhanced visuospatial cognition in orchestral musicians. *Journal of Neuroscience, 27*, 3799-3806.

Spreen, O., & Strauss E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.

Stoesz, B. M., Jakobson, L. J., Kilgour, A. R., & Lewycky, S. T. (2007). Local processing advantage in musicians: Evidence from disembedding and constructional tasks. *Music Perception, 25*, 153-165.

Uttl, B. (2002). North American Adult Reading Test: Age norms, reliability, and validity. *Journal of Clinical and Experimental Neuropsychology, 24*, 1123-1137.

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

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

WIENS, A. N., BRYAN, J. E., & CROSSEN, J. R. (1993). Estimating WAIS-R FSIQ from the National Adult Reading Test-Revised in normal subjects. *The Clinical Neuropsychologist, 7*, 70-84.

WITKIN, H. A., OLTMAN, P. K., RASKIN, E., & KARP, S. A. (1971). *A manual for the Embedded Figures Tests*. Palo Alto, CA: Consulting Psychologists Press.

YONG, B. C. K., & MCBRIDE-CHANG, C. (2007). Emotion perception for faces and music: Is there a link? *Korean Journal of Thinking and Problem Solving, 17*, 57-65.