AMERICAN PSYCHOLOGICAL ASSOCIATION DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

1931-3896/19/$12.00
http://dx.doi.org/10.1037/aca0000263

Correlation = Causation? Music Training, Psychology, and Neuroscience

E. Glenn Schellenberg
University of Toronto

Although students in introductory psychology courses know that it is wrong to infer causation from correlation, scholars sometimes do so with impunity. The present study sought to test the hypothesis that the problem is systematic in studies of music training, and whether it may be attenuated or exacerbated among neuroscientists compared to behavioral psychologists. The sample comprised 114 journal articles published since 2000, each of which examined associations between music training and nonmusical behavior or brain structure and function. Articles were classified as neuroscience or psychology based on the method and/or journal. Raters blind to the question about discipline determined from the titles and abstracts whether the authors made a causal inference. Inferences of causation were common in both disciplines, but the problem was particularly acute among neuroscientists, with their odds of inferring causation more than twice as great as those of psychologists. The results highlight a narrow-minded focus on learning and the environment among researchers who study music training, as well as an apparent disregard of findings from studies of far transfer, behavioral genetics, and other factors that distinguish individuals with or without music training.

Keywords: correlation, transfer, psychology, neuroscience, music training

Supplemental materials: http://dx.doi.org/10.1037/aca0000263.supp

Students in introductory psychology courses are taught that correlation does not imply causation. Although this concept is simple, researchers often infer a causal direction from correlational results, or exhibit a strong but unwarranted interpretive bias. The goal of the present study was to document systematic interpretive biases in research on associations between music training and nonmusical variables, including brain structure and function. An open question was whether the prevalence of inferences of causation from correlational or quasiexperimental data might differ between neuroscientists and behavioral psychologists.

In the case of music training, causal inferences rest on the assumption that individuals with or without music lessons vary only in the extent of their training. This assumption is unfounded because individuals with music training differ from other individuals in multiple ways. For example, among 10- to 12-year-old Canadian children (Corrigall, Schellenberg, & Misura, 2013), duration of music lessons has positive associations with age, two measures of socioeconomic status (family income, parents' education), two measures of general cognitive ability (IQ, average grade in school), and two dimensions of personality (conscientiousness, openness-to-experience). Duration of music training is also correlated positively with duration of involvement in nonmusical out-of-school activities. Although these data are correlational, it is doubtful that music training is the causal agent in each instance.

Moreover, large-scale twin studies document a genetic component to music aptitude, extent of practicing, and musical skill or achievement (Hambrick & Tucker-Drob, 2015; Mosing, Madison, Pedersen, Kuja-Halkola, & Ullén, 2014), which must be instantiated in the brain. In other words, associations between music training and brain structure or function do not inform the issue of causation. The link between music training and general cognitive ability is also influenced by genetics (Mosing, Madison, Pedersen, & Ullén, 2016), as is the musical instrument and genre of choice (Mosing & Ullén, 2018). The genetic contribution to practice is particularly informative, because practice is typically considered to be one of the main environmental contributors to musical achievement (Howe, Davidson, & Sloboda, 1998). Indeed, the contribution of deliberate practice to individual differences in music performance is much smaller than we once thought (Macnamara, Hambrick, & Oswald, 2014), but nevertheless important for individuals who have the genetic potential to become an accomplished musician (Hambrick & Tucker-Drob, 2015).

Claims that music training causes improvements in general cognitive abilities (e.g., executive functions; Okada & Slevc, 2019), language abilities (e.g., speech perception, reading; Tierney & Kraus, 2013), or visual-spatial abilities (Anaya, Pisoni, & Kro-

Funded by the Natural Sciences and Engineering Research Council of Canada. Assisted by Michael Carnovale, Muhammad Khan, and Eric Tu. Susan Ehrlich, Elizabeth Johnson, Mattson Ogg, Sandra Trehub, and Ellen Winner provided helpful comments on earlier drafts.

Correspondence concerning this article should be addressed to E. Glenn Schellenberg, Department of Psychology, University of Toronto, 3359 Mississauga Road, Mississauga, Ontario, Canada, L5L 1C6. E-mail: g.schellenberg@utoronto.ca

nenberger, 2017) are claims of *far transfer*, which is rare or nonexistent across domains (Woodworth & Thorndike, 1901), particularly when a study is well designed—with random assignment and an active control group (Sala & Gobet, 2017a, 2019). The best evidence for the effects of music training on general cognitive abilities comes from a paper published in *Psychological Science*—"Music Lessons Enhance IQ" (Schellenberg, 2004), the results of which cannot be replicated (Haywood et al., 2015) and are belied by a recent meta-analysis that reported null findings (Sala & Gobet, 2017b). In any event, effect sizes in correlational studies of music training and general cognitive ability are comparatively much larger (e.g., 2/3–1 *SD*; Schellenberg, 2011; Schellenberg & Mankarious, 2012) than those found in interventions with random assignment. Thus, in the real world, other factors, such as preexisting differences, must play a role in the observed associations.

To summarize, inferences of causal effects from correlational studies of music training violate the rules of science, and are undermined by: (a) preexisting individual differences in demographics, cognitive ability, and personality; (b) results from studies of behavioral genetics; and (c) the rarity of far-transfer effects. The present investigation sought to document the extent of this problem by collecting a comprehensive sample of journal articles published during the 21st century. Another goal was to test whether the problem might vary in prevalence based on discipline. For example, neuroscience is often considered to be a "harder" science than psychology, presumably because imaging techniques are more objective than behavioral measures, and less susceptible to factors such as response bias and demand characteristics. From this view, the prevalence of incorrect inferences might be lower among neuroscientists than among psychologists. Alternatively, brain images make links between cognition and the brain seem less abstract and more real (McCabe & Castel, 2008), such that slippage from association to causation might be greater among neuroscientists than it is among psychologists. Indeed, although skeptics have criticized researchers who attempt to explain complex behavior at the neuronal level (Legrenzi & Umilta, 2011; Satel & Lilienfeld, 2013), brain data and images may promote the likelihood of concluding that associations with music training are causal phenomena.

The design of the present study was straightforward. The sample comprised a large sample of articles written by authors who used correlational or quasiexperimental designs to study associations between music training and nonmusical variables, including perceptual and cognitive abilities, as well as measures of brain structure and function. Each article was classified as neuroscience or psychology. Two raters (A and B), who were blind to the idea of comparing disciplines, determined independently whether or not the authors of each article inferred causation.

Method

In January to April of 2018, a research assistant searched for all articles published since 2000 that had correlational designs and tested for associations between music training and a nonmusical ability, brain structure, or brain function. In June of 2018, a second research assistant was asked to find additional articles that may have been overlooked. The assistants used multiple search engines (Google Scholar, PsycINFO, Web of Science) and the following search terms: *music training*, *musical training*, *musicians*, *nonmusicians*, *music lessons*, *musical lessons*, *music experience*, and *musical experience*. Inclusion criteria were an original empirical (quantitative) study, a correlational or quasiexperimental (cross-sectional) design, and a dependent variable that measured an ability or anatomical structure not directly related to music (e.g., speech perception, cognitive abilities, brain-stem responses to sound, structural and functional brain measures). Exclusion criteria included random assignment, a null simple association (i.e., no finding to misinterpret), and a longitudinal design. Although longitudinal designs may also be quasiexperimental (and therefore involve problems of self-selection), they nevertheless examine change over time, and the focus here was on the most egregious examples of inferring causation from correlational data. Review articles and very short reports without abstracts were also excluded.

A total of 114 articles met the criteria for inclusion in the final sample (see the online supplemental materials). The vast majority of the articles ($N = 101$) reported results from studies with quasiexperimental (cross-sectional) designs, which compared musicians and nonmusicians, or individuals with or without music training. Some of these quasiexperimental studies ($N = 11$) had three groups: controls and two groups who varied in amount of music training/experience, or in the specific instrument the musicians played. An additional 13 articles had correlational designs, treating duration of music training as a continuous variable. Small sample sizes precluded the possibility of determining whether these relatively subtle differences in design features influenced the results.

Each article was classified according to whether it belonged to neuroscience or psychology (1 or 0, respectively). Articles were classified as neuroscience if the authors used a neuroscientific method (EEG/ERP, MEG, MRI, fMRI, PET, TMI), or if the journal had a title that included *brain*, *neuroscience*, *neuropsychology* or other terms that referred to the brain or physiology (e.g., *cerebral*, *cortex*, *laterality*). General science journals (e.g., *PLoS ONE*, *PNAS*) were not considered to be neuroscience journals. Of the articles assigned to the neuroscience category, most used a neuroscientific method (64 of 71), and most were published in a neuroscience journal (50 of 71). Forty-three articles met both criteria, 21 met only the method criterion, and seven met only the journal criterion. All other articles were classified as psychology.

The second assistant (hereafter Rater A), who was unaware of any interest in differences between disciplines, was asked to determine from the title and/or abstract of each article whether the authors inferred causation (1 = *yes*, 0 = *no*). If the title contained a causal verb (e.g., *enhance*, *affect*, *promote*, *facilitate*, *lead to*, *modulate*, *impact*, *strengthen*, *heighten*, *shape*, *limit*, *shift*, *contribute*), the authors were deemed to infer causation. If the title was unclear in this regard, the assistant went on to read the abstract, searching for the same causal language. If the authors used a causal verb combined with a qualifying term in the same sentence (*may*, *maybe*, *possibly*, or *perhaps*), they were classified as 0 (no causal inference). Moreover, if music training was considered to be the outcome variable rather than a predictor variable, the article was coded 0. Finally, a third assistant (hereafter Rater B), who was also unaware about questions involving discipline, made independent judgments using the same method. Neither rater had worked previously in the author's laboratory, or had read any of his journal articles.

Agreement between Raters A and B was initially modest (76%), primarily because of instances in which authors used hedge verbs (e.g., *suggest*) in addition to causal verbs to advance the position that music lessons cause nonmusical outcomes. Nevertheless, after the two raters—still blind to the question about discipline—met to discuss and reconcile discrepancies, they reached almost complete agreement (96%).

Results

Causation was operationally defined in five different way to ensure that the results were not specific to a single definition. Specifically, separate analyses were conducted using the original ratings from Rater A, the original ratings from Rater B, instances in which Raters A and B initially agreed, instances of agreement after discussion, and the sum of the two original ratings (0 = *no causation*, 1 = *disagreement*, 2 = *causation*).

For each of the five methods, a chi-square test of independence examined whether the likelihood of inferring causation differed across disciplines. The results, illustrated in Figure 1, were identical across methods in terms of statistical significance, but effect sizes and exact *p* values varied. As shown in the figure, most of the 114 articles in the sample (62%) were written by neuroscientists, which simply highlights the field's interest in music training. As

Rater A

	Psychology	Neuroscience	
No Causation	22	19	41
Causation	21	52	73
	43	71	114

χ^2	d.f.	*p*	O.R.	ϕ
6.92	1	.009	2.87	.246

Rater B

	Psychology	Neuroscience	
No Causation	24	20	44
Causation	19	51	70
	43	71	114

χ^2	d.f.	*p*	O.R.	ϕ
8.64	1	.003	3.22	.275

Articles with initial agreement between raters

	Psychology	Neuroscience	
No Causation	19	10	29
Causation	16	42	58
	35	52	87

χ^2	d.f.	*p*	O.R.	ϕ
11.57	1	<.001	4.99	.365

Articles with final agreement between raters

	Psychology	Neuroscience	
No Causation	21	16	37
Causation	19	53	72
	40	69	109

χ^2	d.f.	*p*	O.R.	ϕ
9.70	1	.002	3.66	.298

Initial agreement summed

	Psychology	Neuroscience	
No Causation	19	10	29
Disagreement	8	19	27
Causation	16	42	58
	43	71	114

χ^2	d.f.	*p*	Cramer's V
12.83	2	.002	.335

Figure 1. Results from tests of independence between discipline and whether articles inferred causation from correlational data. *OR* = odds ratio. φ and Cramer's V are measures of association (0 = *no association*, 1 = *perfect association*). Each analysis revealed that inferences of causation from correlational data were frequent among psychologists and neuroscientists, but more so among neuroscientists.

expected, in a substantial number of articles (from 61% to 67%), the authors inferred causation. Across analyses, neuroscientists were significantly more likely than psychologists to do so. Odds ratios (calculable for 2 × 2 designs) varied from 2.9 to 5.0. In other words, the odds of inferring causation from correlation were consistently at least 2.9 times greater among neuroscientists than among psychologists. Psychologists made causal inferences a maximum of 49% of the time, whereas a clear majority of neuroscience articles made such mistakes, reaching a maximum of 81% in one instance (for articles with initial agreement between raters).

An impact factor from 2017 was assigned to each article based on data published by Thomson-Reuters/Clarivate Analytics (*InCites Journal Citation Reports*). Three of 114 articles were published in journals with no available impact factor and excluded from further analysis. Impact factors varied more for neuroscience than for psychology, $p = .002$ (Levene's test), which motivated use of an unequal-variances *t* test to compare means. On average, impact factors were higher for neuroscience articles ($M = 4.47$, $SD = 3.77$) than for psychology articles ($M = 1.90$, $SD = 0.93$), $t(82.52) = 5.41$, $p < .001$, which raised the counterintuitive possibility that as the quality (i.e., impact factor) of a journal increases, it also becomes more likely to publish articles written by authors who infer causation from correlation. Nevertheless, impact factor was independent of inferring causation across the five methods, *ps* > .3. Moreover, when logistic regression was used to predict inferences of causation as a joint function of discipline and impact factor, impact factor did not add predictive power in any instance, *ps* > .3. Neuroscientists continued to infer causation more than psychologists, however, even with impact factor held constant, *ps* < .02.

Discussion

As predicted, inferences of causation from correlational designs were frequent in studies of associations between music training and nonmusical abilities, and in studies of associations between music training and brain structure or function. Another major finding was that the likelihood of making such erroneous inferences was greater for neuroscientists than for psychologists.

Are these conclusions valid? There is no baseline comparison rate for determining that inferences of causation from correlation are abnormally frequent, because even a single instance represents an error in scientific logic. Moreover, some articles that met the inclusion criteria are likely to have been overlooked, yet there is no reason to expect that they would differ systematically from those that were included. (The goal was to recruit the population.) The process of determining whether authors inferred causation also had gray areas, but the results did not differ when only instances of original or final interrater agreement were considered, or when instances of disagreement were considered as a separate category. Finally, some authors may have hedged their inferences later in the text, but this fact does not ameliorate the causal inference in the title and/or abstract, which would influence readers' overall interpretation of the article (van Dijk & Kintsch, 1983).

The problem of unjustified causal inferences is likely to stem from current fascination with changes in brain and behavior that occur as a consequence of learning and the environment. Indeed, *plasticity* has become an industry of its own, as evidenced by best-selling trade books (Doidge, 2007; Hurley, 2014) and commercial brain-training programs (e.g., Lumosity), in addition to large grants for research programs on plasticity. Despite the overwhelming evidence that behavioral and cognitive development is the consequence of nature and nurture (e.g., Deary, 2012; Långström, Rahman, Carlström, & Lichtenstein, 2010; Schellenberg, 2015; Weikum et al., 2013), many studies that examine music training as a predictor of nonmusical abilities appear to do so in the vein of what Detterman, Gabriel, and Ruthsatz (1998) call "absurd environmentalism." That is, they focus on environmental factors while ignoring the literature on the heritability of musical ability. Academic departmentalization may exacerbate this problem inadvertently by allowing disciplines (e.g., neuroscience and genetics) to exist in parallel universes. Brain data and images are also known to reify correlational results (McCabe & Castel, 2008; Racine, Bar-Ilan, & Illes, 2005), with the likely consequence that interpretations from neuroscience receive less scrutiny than those from psychology, not only by the general public and policymakers, but also by reviewers, editors, and funding agencies.

Are inferences of causation from correlational data limited to research on music training? Probably not, although it is difficult to find other research areas where such logical errors are as frequent, even though there are many other forms of experience that correlate with cognitive abilities or neural characteristics. For example, playing chess was once thought to be an ideal model for the study of expertise. Indeed, Simon and Chase (1973) claimed that chess was the *Drosophilia* for the study of artificial intelligence and the accumulation of knowledge more generally. Chess experience does indeed predict chess expertise (Bilalić, McLeod, & Gobet, 2007; Grabner, Stern, & Neubauer, 2007) and remains useful as a model in this regard. Chess players also tend to have above-average intelligence (Bilalić et al., 2007; Grabner et al., 2007) and to be predominantly male (Chabris & Glickman, 2006), which suggests that, as with learning to play music, self-selection plays a role in who chooses to play chess. Despite the association with general cognitive ability, evidence that playing chess causes far-transfer effects is minimal (Sala & Gobet, 2016).

Other mentally demanding activities such as crossword puzzles are associated with enhanced cognition in the elderly, but these advantages are evident earlier in life, with similar age-related declines (Salthouse, 2006). Speaking a second language appears to offer another interesting parallel with taking music lessons (Bialystok & DePape, 2009). In contrast to music training, however, self-selection plays a much smaller role when individuals learn the language that is spoken in their environment (i.e., forced bilingualism). Moreover, evidence that bilingualism confers cognitive advantages (Bialystok, 2015) remains controversial (Valian, 2015), and is undermined by a bias favoring the publication of positive results (de Bruin, Treccani, & Della Sala, 2015), as well as by findings from a recent meta-analysis (Lehtonen et al., 2018).

Other possible parallels to music training include commercial brain-training programs, laboratory-based working memory training, and video games. Brain-training programs improve performance on the trained task, which sometimes extends to small improvements on similar tasks, but performance does not transfer to distantly related cognitive tasks or to cognition in everyday life (Simons et al., 2016). Although working memory training in the laboratory is claimed to have far-transfer effects that extend to fluid intelligence (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008), there is little evidence that such effects are long-lasting except, perhaps, for certain individuals (Jaeggi, Buschkuehl, Jonides, & Shah, 2011). Furthermore, meta-analyses raise serious doubts about whether such effects are even reliable (Melby-Lervåg & Hulme, 2013; Sala & Gobet, 2017c).

Perhaps the closest parallel with music training comes from studies of individuals who play video games. Action video-games appear to cause improvements in low-level visual perception (Green & Bavelier, 2007; Li, Polat, Makous, & Bavelier, 2009) and attention (Green & Bavelier, 2003), which are arguably near- rather than far-transfer effects. Another difference with music training is that psychologists and neuroscientists who study video games tend to be meticulous about distinguishing correlational results from those of interventions with random assignment (Green, Gorman, & Bavelier, 2016), possibly because earlier findings suggested that violent video games were a risk factor for negative social behaviors (Anderson et al., 2010). In any case, positive transfer effects of playing video games are highly circumscribed. When exposure to video games (action and otherwise) is examined more generally, associations with cognitive abilities disappear (Sala, Tatlidil, & Gobet, 2018; Unsworth et al., 2015). In short, video-game research has yielded positive, negative, and null results.

Epidemiological studies represent another area in which documented associations from large samples of individuals are often interpreted as having a causal direction, such that a predictor variable is viewed as causally altering the chances of contracting a specific disease or health problem. For example, a meta-analysis of studies that included almost 2 million participants concluded that moderate alcohol consumption reduces some adverse cardiovascular outcomes, whereas heavy drinking increases others (Bell et al., 2017). Inferences from epidemiological studies are different from those made in the music-training literature because: (a) medical research has a history of considering extraneous variables (Greenland, Pearl, & Robins, 1999; Schisterman, Cole, & Platt, 2009; Smith & Hemani, 2014), and (b) the reverse causal direction is often implausible or impossible because of the timeline. By contrast, there is an abundance of evidence suggesting that multiple genetic and environmental individual differences influence who takes music lessons, particularly for extended durations of time. Music training may indeed serve to amplify such differences, but characteristics of the particular pedagogy, the teacher, the student, and the interaction between teacher and student mean that these are likely to be idiosyncratic rather than systematic (Schellenberg, 2019).

In studies of music training, frequent inferences of causation from correlation appear to arise from a unique combination of converging factors, which serve to set it apart from other research domains. One such factor is that long-term interventions with random assignment are virtually impossible to conduct because of artificiality and attrition (Schellenberg, 2004; Slater et al., 2015). Another is that music training has multiple associations with positive variables, but no associations with negative variables (for review see Schellenberg & Weiss, 2013). Psychologists and neuroscientists working in the area also exhibit naiveté about the (im)possibility of far-transfer effects, unbridled enthusiasm for plasticity, disregard of findings from behavioral genetics, and unwarranted faith in results that include brain data and images. Finally, the idea of positive cognitive and neural side effects from music training (and other pleasurable activities) is inherently appealing, despite the fact that there is little positive evidence in this regard. It is a disservice, however, to offer false hope, wittingly or otherwise, to the public, educators, and other researchers.

References

Anaya, E. M., Pisoni, D. B., & Kronenberger, W. G. (2017). Visual-spatial sequence learning and memory in trained musicians. *Psychology of Music, 45,* 5–21. http://dx.doi.org/10.1177/0305735616638942

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., . . . Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin, 136,* 151–173. http://dx.doi.org/10.1037/a0018251

Bell, S., Daskalopoulou, M., Rapsomaniki, E., George, J., Britton, A., Bobak, M., . . . Hemingway, H. (2017). Association between clinically recorded alcohol consumption and initial presentation of 12 cardiovascular diseases: Population based cohort study using linked health records. *British Medical Journal, 356,* j909. http://dx.doi.org/10.1136/bmj.j909

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives, 9,* 117–121. http://dx.doi.org/10.1111/cdep.12116

Bialystok, E., & Depape, A.-M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 565–574. http://dx.doi.org/10.1037/a0012735

Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence?—A study with young chess players. *Intelligence, 35,* 457–470. http://dx.doi.org/10.1016/j.intell.2006.09.005

Chabris, C. F., & Glickman, M. E. (2006). Sex differences in intellectual performance: Analysis of a large cohort of competitive chess players. *Psychological Science, 17,* 1040–1046. http://dx.doi.org/10.1111/j.1467-9280.2006.01828.x

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology, 4,* 222. http://dx.doi.org/10.3389/fpsyg.2013.00222

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology, 63,* 453–482. http://dx.doi.org/10.1146/annurev-psych-120710-100353

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26,* 99–107. http://dx.doi.org/10.1177/0956797614557866

Detterman, D. K., Gabriel, L. T., & Ruthsatz, J. M. (1998). Absurd environmentalism. *Behavioral and Brain Sciences, 21,* 411–412. http://dx.doi.org/10.1017/S0140525X98271238

Doidge, N. (2007). *The brain that changes itself*. New York, NY: Viking.

Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica, 124,* 398–420. http://dx.doi.org/10.1016/j.actpsy.2006.07.008

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534–537. http://dx.doi.org/10.1038/nature01647

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18,* 88–94. http://dx.doi.org/10.1111/j.1467-9280.2007.01853.x

Green, C. S., Gorman, T., & Bavelier, D. (2016). Action video-game training and its effects on perception and attentional control. In T. Strobach & J. Karbach (Eds.), *Cognitive training: An overview of features and applications* (pp. 107–116). Basel, Switzerland: Springer International. http://dx.doi.org/10.1007/978-3-319-42662-4_10

Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology, 10,* 37–48. http://dx.doi.org/10.1097/00001648-199901000-00008

Hambrick, D. Z., & Tucker-Drob, E. M. (2015). The genetics of music accomplishment: Evidence for gene-environment correlation and interaction. *Psychonomic Bulletin & Review, 22,* 112–120. http://dx.doi.org/10.3758/s13423-014-0671-9

Haywood, S., Griggs, J., Lloyd, C., Morris, S., Kiss, Z., & Skipp, A. (2015). *Creative futures: Act, sing, play. Evaluation report and executive summary*. London, UK: Educational Endowment Foundation. Retrieved from https://files.eric.ed.gov/fulltext/ED581247.pdf

Howe, M. J., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences, 21,* 399–407. http://dx.doi.org/10.1017/S0140525X9800123X

Hurley, D. (2014). *Smarter: The new science of building brain power*. New York, NY: Plume.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 6829–6833. http://dx.doi.org/10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 10081–10086. http://dx.doi.org/10.1073/pnas.1103228108

Långström, N., Rahman, Q., Carlström, E., & Lichtenstein, P. (2010). Genetic and environmental effects on same-sex sexual behavior: A population study of twins in Sweden. *Archives of Sexual Behavior, 39,* 75–80. http://dx.doi.org/10.1007/s10508-008-9386-1

Legrenzi, P., & Umilta, C. (2011). *Neuromania: On the limits of brain science*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199591343.001.0001

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144,* 394–425. http://dx.doi.org/10.1037/bul0000142

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience, 12,* 549–551. http://dx.doi.org/10.1038/nn.2296

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25,* 1608–1618. http://dx.doi.org/10.1177/0956797614535810

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107,* 343–352. http://dx.doi.org/10.1016/j.cognition.2007.07.017

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49,* 270–291. http://dx.doi.org/10.1037/a0028228

Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., & Ullén, F. (2014). Practice does not make perfect: No causal effect of music practice on music ability. *Psychological Science, 25,* 1795–1803. http://dx.doi.org/10.1177/0956797614541990

Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén, F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science, 19,* 504–512. http://dx.doi.org/10.1111/desc.12306

Mosing, M. A., & Ullén, F. (2018). Genetic influences on musical specialization: A twin study on choice of instrument and music genre. *Annals of the New York Academy of Sciences, 1423,* 427–434. http://dx.doi.org/10.1111/nyas.13626

Okada, B. M., & Slevc, L. R. (2019). Musical training: Contributions to executive function. In M. Bunting, J. Novick, M. Dougherty, & R. W. Engle (Eds.), *An integrative approach to cognitive and working memory training: Perspectives from psychology, neuroscience, and human development*. New York, NY: Oxford University Press. http://dx.doi.org/10.13016/M2GM81P70

Racine, E., Bar-Ilan, O., & Illes, J. (2005). fMRI in the public eye. *Nature Reviews Neuroscience, 6,* 159–164. http://dx.doi.org/10.1038/nrn1609

Sala, G., & Gobet, F. (2016). Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis. *Educational Research Review, 18,* 46–57. http://dx.doi.org/10.1016/j.edurev.2016.02.002

Sala, G., & Gobet, F. (2017a). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26,* 515–520. http://dx.doi.org/10.1177/0963721417712760

Sala, G., & Gobet, F. (2017b). When the music's over. Does musical skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review, 20,* 55–67. http://dx.doi.org/10.1016/j.edurev.2016.11.005

Sala, G., & Gobet, F. (2017c). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53,* 671–685. http://dx.doi.org/10.1037/dev0000265

Sala, G., & Gobet, F. (2019). Cognitive training does not enhance general cognition. *Trends in Cognitive Sciences, 23,* 9–20. http://dx.doi.org/10.1016/j.tics.2018.10.004

Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin, 144,* 111–139. http://dx.doi.org/10.1037/bul0000139

Salthouse, T. A. (2006). Mental exercise and mental aging: Evaluating the validity of the "use it or lose it" hypothesis. *Perspectives on Psychological Science, 1,* 68–87. http://dx.doi.org/10.1111/j.1745-6916.2006.00005.x

Satel, S., & Lilienfeld, S. O. (2013). *Brainwashed: The seductive appeal of mindless neuroscience*. New York, NY: Basic.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15,* 511–514. http://dx.doi.org/10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology, 102,* 283–302. http://dx.doi.org/10.1111/j.2044-8295.2010.02000.x

Schellenberg, E. G. (2015). Music training and speech perception: A gene-environment interaction. *Annals of the New York Academy of Sciences, 1337,* 170–177. http://dx.doi.org/10.1111/nyas.12627

Schellenberg, E. G. (2019). Music training, music aptitude, and speech perception. *Proceedings of the National Academy of Sciences of the United States of America, 116,* 2783–2784. http://dx.doi.org/10.1073/pnas.1821109116

Schellenberg, E. G., & Mankarious, M. (2012). Music training and emotion comprehension in childhood. *Emotion, 12,* 887–891. http://dx.doi.org/10.1037/a0027971

Schellenberg, E. G., & Weiss, M. W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 499–550). Amsterdam, the Netherlands: Elsevier. http://dx.doi.org/10.1016/B978-0-12-381460-9.00012-2

Schisterman, E. F., Cole, S. R., & Platt, R. W. (2009). Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology, 20,* 488–495. http://dx.doi.org/10.1097/EDE.0b013e3181a819a1

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61,* 394–403.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17,* 103–186. http://dx.doi.org/10.1177/1529100616661983

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research, 291,* 244–252. http://dx.doi.org/10.1016/j.bbr.2015.05.026

Smith, G. D., & Hemani, G. (2014). Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics, 23,* R89–R98. http://dx.doi.org/10.1093/hmg/ddu328

Tierney, A., & Kraus, N. (2013). Music training for the development of reading skills. *Progress in Brain Research, 207,* 209–241. http://dx.doi.org/10.1016/B978-0-444-63327-9.00008-4

Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science, 26,* 759–774. http://dx.doi.org/10.1177/0956797615570367

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18,* 3–24. http://dx.doi.org/10.1017/S1366728914000522

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Weikum, W. M., Brain, U., Chau, C. M. Y., Grunau, R. E., Boyce, W. T., Diamond, A., & Oberlander, T. F. (2013). Prenatal serotonin reuptake inhibitor (SRI) antidepressant exposure and serotonin transporter promoter genotype (SLC6A4) influence executive functions at 6 years of age. *Frontiers in Cellular Neuroscience, 7,* 180. http://dx.doi.org/10.3389/fncel.2013.00180

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions (I). *Psychological Review, 8,* 247–261. http://dx.doi.org/10.1037/h0074898

Received January 16, 2019
Revision received March 19, 2019
Accepted April 24, 2019 ■