ABSTRACT

Title of Thesis:	MUSICAL TRAINING AND EXECUTIVE FUNCTIONS
	Brooke M. Okada, Master of Science, 2016
Thesis Directed By:	Dr. L. Robert Slevc Department of Psychology

Learning and performing music draws on a host of cognitive abilities. One likely aspect of cognition that may be related to musical training is executive function. To date, many studies have investigated this relationship; however, results from such studies are mixed and difficult to compare. In part, this is because most studies look at only one specific cognitive process, and even studies looking at the same process use different experimental tasks. The current study addresses these issues by administering a well-validated EF battery of multiple tasks tapping each EF component (Friedman et al., 2008) and a comprehensive measure of musical training (Müllensiefen, Gingras, Musil, & Stewart, 2014) to obtain reliable measures of individual differences in EF and musical experience. Results suggest that there is positive relationship between musical training and performance on updating tasks, but this relationship is not observed with performance on inhibition or shifting tasks.

MUSICAL TRAINING AND EXECUTIVE FUNCTIONS

by

Brooke M. Okada

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2016

Advisory Committee:
Professor L. Robert Slevc, Chair
Professor Michael Dougherty
Professor Kenneth Elpus

Table of Contents

Introduction

There has been growing interest in music lessons (musical training) and their associated non-musical benefits (see, e.g., the special edition of the journal *Music Perception* on "Music Training and Nonmusical Abilities" Schellenberg & Winner, 2011). The National Association for Music Education posits links between children who take music lessons and higher academic achievement as well as "creativity, curiosity, and personal motivation" in advocating for music education's place in schools (Broader Minded Media Resources, 2015). Thus, the research investigating transfer effects of musical training spans a wide range of domains, including social benefits (e.g., joint music making leads to pro-social behavior (Kirschner & Tomasello, 2010)) as well as benefits in the cognitive domain (see Benz, Sellaro, Hommell, & Colzato, 2015, for a review).

This body of work investigating the transfer effects of musical training fits into the broader literature of transfer effects of cognitive training, with the idea that repetitious training on a cognitively demanding task (e.g., computer-based N-back training as in Jaeggi, Buschkuehl, Jonides, & Perrig, 2008) may transfer to and increase performance on similar tasks (near transfer, e.g., other working memory tasks) or even to more distantly related cognitive abilities (far transfer, e.g., intelligence). It is important to note that the overall evidence for the body of literature concerning working memory training and transfer is mixed (Au et al., 2015; Dougherty, Hamovitz, & Tidwell, 2015; Melby-Lervåg & Hulme, 2013). However, one major limitation of most cognitive training regimens is that they are administered during hours of repetitive and tedious computer tasks. In contrast, musical training is a more enjoyable and engaging task that is regularly pursued for long periods of time. Thus, musical training offers several potential advantages: it might be less susceptible to attrition and subjects may be more intrinsically motivated since they are receiving the direct benefit of learning to play an instrument.

Even if musical training does not lead to far transfer to more general cognitive abilities (cf. mixed findings from other cognitive training tasks; e.g., Melby-Lervåg & Hulme, 2013), investigating a possible link between musical training and non-musical cognitive abilities is still a worthwhile endeavor for a few reasons. Finding a positive association between musical training and cognitive abilities could be evidence of transfer effects, or it could reflect selection bias, predispositions in cognitive abilities that influence who takes music lessons (Elpus, 2013). If the latter is the case, this may be evidence of music training exaggerating these pre-existing individual differences (i.e., a gene-environment interaction (Schellenberg, 2015)), and may shed light on what types of cognitive abilities are relevant to music learning, thus helping us better understand the ways music interacts with general cognitive abilities.

To date, many studies have investigated various cognitive processes associated with musical training. Taking music lessons has been found to correlate with multiple cognitive processes including, but not limited to: full-scale IQ and academic ability (Schellenberg, 2006; but see Schellenberg & Moreno, 2010; Schellenberg, 2015); selective auditory attention (Strait & Kraus, 2011); verbal short term memory (Chan, Ho, & Cheung, 1998; Ho, Cheung, & Chan, 2003; Jakobson, Cuddy, & Kilgour, 2003; Jakobson, Lewycky, Kilgour, & Stoesz, 2008); reading ability (Butzlaff, 2000; Lamb & Gregory, 1993); mathematical ability (Vaughn, 2000); spatial skills (Bilhartz, Bruhn, & Olson, 2000); processing speed (Bugos & Mostafa, 2011); second language learning (Slevc & Miyake, 2006); set shifting (Degé, Kubicek, & Schwarzer, 2011; Hanna-Pladdy & MacKay, 2011); inhibition (Bialystok & DePape, 2006; Degé et al., 2011); and working memory updating (Oechslin, Van De Ville, Lazeyras, Hauert, & James, 2013; Pallesen et al., 2010; Slevc, Davey, Buschkuehl, & Jaeggi, 2016).

One hurdle in interpreting the literature on the cognitive benefits of music training is that each study typically looks at only one specific cognitive process. For example, Moreno et al. (2011) look only at inhibition and Chan et al. (1998) look only at verbal short term memory. Even studies looking at the same processes use different tasks to measure that process (e.g., inhibition measured with a Stroop task (Schellenberg, 2011) or a go/no-go task (Moreno et al., 2011)) with mixed results. Furthermore, many studies use only one task to measure a given process, which is potentially problematic. Any complex task may be measuring the construct of interest properly, but will also be tapping other types of processes as well (i.e., the "task impurity problem," Miyake et al., 2000), so the extent to which any one task is accurately measuring and reflecting the intended construct is unknown. For example, a go/no-go task (where one needs to inhibit responding when given a cue) certainly involves inhibitory ability, but presumably also requires processes such as sustained attention and the memory for maintenance of task instructions and goals. Thus, even with the abundance of studies on musical training, it is not yet clear which specific cognitive abilities are related to musical training.

An additional complication is that previously documented relationships between musical training and cognitive abilities might actually reflect some other underlying ability. Specifically, Schellenberg (2004; 2006; 2011) suggested that there may be a more general cognitive capacity (i.e., general intelligence (IQ)) that underlies the links between musical training and these various cognitive processes. Consistent with this hypothesis, he found that full-scale IQ correlated with amount of musical training (Schellenberg, 2006; 2011), and the same results were also found experimentally: children randomly assigned to music lessons (keyboard or Kodály voice lessons) had greater gains in full-scale IQ than those taking drama lessons or no lessons (Schellenberg, 2004). He concluded that this could account for the different findings because increases in general intelligence (as assessed by the WISC-III, which contains subtests tapping various cognitive processes (i.e., Verbal Comprehension, Perceptual Organization, Freedom from Distractibility, and Processing Speed)), may be the source of the increased benefits seen in other cognitive processes.

However, IQ has been found to be stable across the lifespan (Deary, 2001), and Schellenberg (2011) remarks that this “makes it unlikely that a single environmental factor [music training] could have much of an impact” (Schellenberg, 2011, p. 287). This suggests that musical training / cognitive ability relationships might instead reflect individual differences in who decides to take music lessons (Elpus, 2013; Schellenberg, 2015). Therefore, other general cognitive processes that may be more modifiable should be examined, as they might be underlying music training's link with the various specific cognitive abilities mentioned earlier. One such cognitive process that has been theorized to be associated with musical training is executive function (e.g., Hannon & Trainor, 2007; Moreno & Farzan, 2015; Okada & Slevc, in press).

Executive Functions

Executive functions (EFs), also known as cognitive control, are a set of top-down processes that regulate behavior and cognition according to task demands (Diamond, 2013; Miyake & Friedman, 2012; Novick, Kan, Trueswell, & Thompson-Schill, 2009). Most models of executive functions postulate three related, but separable components: inhibition (or inhibitory control), shifting (or cognitive flexibility or switching), and updating (or working memory updating) (Diamond, 2013; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Logue & Gould, 2014; Miyake et al., 2000; Miyake & Friedman, 2012). Inhibition requires one to override a prepotent response, shifting requires one to switch between task demands, and updating requires one to constantly add and delete items from working memory (Miyake et al., 2000).

These processes develop through adolescence and are important because they have been associated with quality of life and are predictors of school success (Diamond, 2013). They have also been shown to improve with practice (Diamond, Barnett, Thomas, & Munro, 2007; for a review, see Diamond, 2013), and so there is a growing body of research investigating predictors of EF abilities and how one might improve these EF abilities.

Musical Training and Executive Functions

There have been both correlational studies investigating whether specific executive functions are associated with musical training (for a summary, see Table 1) as well as experimental studies investigating whether musical training may improve specific executive functions (for a summary, see Table 2).

Correlational Evidence

Various findings suggest that **inhibitory control** ability is associated with musical training. These tasks typically involve overriding (or inhibiting) a prepotent response to correctly complete the task goals. For example, Bialystok and DePape (2009) found that musicians were faster than non-musicians at incongruent trials of a spatial Simon arrows task (Simon & Rudell, 1967), in which participants indicated the direction of an arrow (pointing left or right) while ignoring the arrow's location (on the left or right side of the screen). Musicians also performed better on a pitch-based auditory Stroop task (Hamers & Lambert, 1972), in which they indicated whether the sung words "high" or "low" were either high or low in pitch. Bialystok and DePape (2009) also report that musicians and non-musicians did not differ on any background cognitive abilities (i.e., fluid intelligence, forward and backward span, and the Trail Making Task), which is somewhat surprising given that Part B of the Trail Making Task is thought to measure shifting ability and has been found to correlate with musical training in other studies (Hanna-Pladdy & MacKay, 2011; Zuk et al., 2014)). Zuk et al. (2014) also report finding no difference between musicians and non-musicians (in two groups of different ages) in a Color-Word interference test, in which inhibition must be used to correctly name the color of the word. In sum,

Study	**Bialystok & DePape, 2009**	**Degé et al., 2011**	**Schellenberg, 2011**	**Slevc et al., 2015**	**Pallesen et al., 2010**	**Oechslin et al., 2013**	**Hanna-Pladdy & MacKay, 2011**	**Zuk et al., 2014 - Children**	**Zuk et al., 2014 - Adults**
Notes on Main / Ancillary Effects	Musicianshad faster RTs on Spatial Simon arrows task and Auditory Stroop task.	Months of lessons correlated with set shifting, inhibition, selective attention, planning, and fluency.	Musicians were only better than non-musicians on Digit Span.	Music ability predicted performance on auditory and visual n-back tasks.	Musicians had faster RTs on n-back and more positive correlation between 2-back performance and BOLD signal	Expert musicians had more accurate performance than non-musicians on visual 3-back.	High activity musicians faster than non-musicians on TMT-B.	Musicians better at Trail Making Test (Shifting).	Musicians better at backward digit span.
EF Correlated with Musical Training? Effect Size?	**Yes** (Simon, partial η2 =.15; Stroop, partial η2 =.09)	**Yes**	**No**	**Yes** (only WM Updating)	**Yes**	**Yes**	**Yes** (TMT-B, η2 =.131) **No** (Letter-Number Sequencing)	**Yes** (TMT-B) **No** (Color-Word Interference, Digit Span Backwards)	**Yes** (Digit Span Backwards) **No** (Color-Word Interference, TMT-B)
EF Component Tested	Inhibition	Inhibition Shifting	Inhibition Shifting	WM Updating Inhibition Shifting	WM Updating	WM Updating	Shifting	WM Updating Inhibition Shifting	WM Updating Inhibition Shifting
EF Measures Assessed	Spatial Simon Arrows task Auditory Stroop task	Attention and EF domain of NEPSY II: Set Shifting Selective Attention Planning Inhibition Fluency	Working Memory (Digit Span) Phonological Fluency Mental Flexibility (WCST) Inhibition (Sun-Moon Stroop) Planning (Tower of Hanoi)	Auditory Stroop task Simon Arrows Auditory 2-, 3-, 4-back Visual n-back letter task Auditory switching task Visual switching task	Auditory 1-back and 2-back	Visual n-back letter task	Letter-Number Sequencing Trail Making Test - Part B	Delis-Kaplan Executive Function System: Trail Making Color-Word Interference Verbal Fluency	Delis-Kaplan Executive Function System: Trail Making Color-Word Interference Verbal Fluency Design Fluency
Additional Measures *Covariate Control?*	Language and Musical Background Questionnaire *Cattell Culture Fair Intelligence Test* *Spatial Span Subtest of Wechsler Memory Scale-III* *Trail Making Test*	*Fluid Intelligence: Culture Fair Test*	*Crystallized Intelligence: Wechsler Abbreviated Scale of Intelligence*	*Age* *SES* *Handedness* *Bilingualism*	Passive listening of major, minor, and dissonant chords (fMRI)	Listening to music syntax violations Advanced Measures of Music Audiation (Gordon, 1989) *Raven's Advanced Progressive Matrices (Raven et al. 2003)*	Verbal intelligence Memory Attention WM Language Functions	Digit Span Backwards and Coding subtests (WISC-IV) *Kaufman Brief Intelligence Test (nonverbal IQ)*	Digit Span Backwards and Coding subtests (WAIS-IV) *WASI (nonverbal and verbal IQ)*
Musician	Must have studied lessons for at least half of his/her life Instrumentalists: played at least 1 of 13 instruments (piano, trumpet, violin, clarinet, guitar, drums, baritone, flute, xylophone, saxophone, cello, organ, and viola) Vocalists: classically trained	Continuous measure: months of instruction on an instrument	9- and 10- year olds: 2+ years of music lessons taken outside school 11- and 12- year olds: 3+ years of music lessons taken outside school	Continuous measure: composite measure of the *Musical Ear Test* and *Ollen Musical Sophistication Index*	Classical musicians who were current students or graudates of Sibelius Music Academy (Finland)	Experts: professional pianists Amateurs: amateur pianists, still musically active	High activity musician: played instrument for 10+ years on a regular basis with formal training Low activity musician: played instrument for 1-9 years with some formal training	At least 2 years of private lessons (mean was 5.2 years of lessons)	Professional musicians who either had a degree in music performance or were currently pursuing one
Non-Musician	Monolinguals who do not have specialized music experience	N/A	No music training outside of what was learned in school	N/A	Minimal music training, including obligatory primary school education	No extracurricular musical education	No formal instrumental music training, could not play instrument, could not read music	No lessons outside of what was learned in school	No lessons outside what was learned in school
Sample Size	22 instrumentalists & 25 vocalists / 24 monolinguals	90	106	96	11 musicians / 10 non-musicians	20 expert musicians, 20 amateur musicians / 19 non-musicians	22 high activity musicians, 27 low activity musicians / 21 non-musicians	15 musicians / 12 non-musicians	15 musicians / 15 non-musicians
Age of Subjects (SD) Music Group / Control Group	25(3.2) years & 24.2(5.1) / 23.5(3.8)	137(10) months	9-12 years	20.84 (3.29) years	28 years / 25 years	24.5(4.5) years, 22.2(3.1) / 24(4.5)	70.8(6.3) years, 69.5(6.6) / 69.7(7.9)	10.9(1.2) years	24.8(3.5) years

Table 1. Summary of correlational evidence of Musical Training and EFs.

there is mixed evidence that musical training is associated with increased inhibitory control ability.

In contrast with Bialystok and DePape (2009), other studies have found that musicians were faster at the Trail Making Test, Part B, where subjects were required to connect dots **shifting** between sequential numbers and letters in alphabetical order (i.e., A, 1, B, 2, C, 3, etc.) (Hanna-Pladdy & MacKay, 2011; Zuk et al., 2014). This was found comparing older adult musicians and non-musicians (Hanna-Paddy & MacKay, 2011), and for children (Zuk et al., 2014). Interestingly, Zuk et al. (2014) did not find this relationship when looking at adult musicians and non-musicians.

Musicians have also been shown to outperform non-musicians on n-back tasks (Oechslin et al., 2013; Pallesen et al., 2010; Slevc et al., 2016), a measure requiring **working memory updating**. In one study, auditory and visual n-back tasks were administered: in the auditory 2-, 3-, and 4-back conditions, participants were required to respond when they heard the same pitch (from a C-Major scale) 2, 3, or 4 notes previously, and in the visual version, participants were asked to respond when they saw the same letter appear 2, 3, or 4 letters previously (Slevc et al., 2016). Music ability predicted n-back performance on both versions, but did not predict performance on auditory or visual versions of inhibition or shifting tasks (Slevc et al., 2016).

In an fMRI study, musicians also had faster reaction times (on the 1-back and 2-back) and were significantly more accurate (on the 1-back only) than non-musicians in another version of auditory n-back, in which participants were presented with a sequence of pitches and were required to report whether the note they heard 1 or 2 notes previously was in the same octave (Pallesen et al., 2010). Although these results do not strongly suggest behavioral benefits, musicians also showed enhanced BOLD responses as a function of N-back difficulty in areas associated with cognitive control (e.g., prefrontal cortex (Miller & Cohen, 2001)). More evidence comes from a study showing that expert pianists were more accurate at a 3-back letter task than non-musicians (Oechslin et al., 2013). However, mixed results from Zuk et al. (2014) show that adult (but not children) musicians showed better performance than non-musicians on the Digit Span Backwards, in which numbers must be held in working memory and correctly manipulated to recall them in backwards order (Zuk et al., 2014).

Thus, advantages for musicians have been shown across all three components of EF tasks. As noted above, because of the potential malleability of EF (Diamond, 2013), this raises the possibility that underlying EF advantages mediate the previous effects linking music lessons and general intelligence (Schellenberg, 2004; 2006). Two of the more comprehensive studies of musical training and EFs addressed this question by assessing multiple EF components (albeit still only with individual tasks) and full-scale IQ in large groups of 9-12 year olds. Schellenberg (2011) found that musical training was associated with full-scale IQ, but found no link between musical training and four of the five EF measures assessed: Phonological Fluency (i.e., naming animals that start with “S” in one minute), Sun-Moon Stroop (i.e., saying “sun” when seeing a picture of a moon and vice versa), Tower of Hanoi (i.e., rearranging rings in order from biggest to smallest on three rods), and Wisconsin Card Sorting Test (i.e., categorizing cards by shape or color), but there was an effect for Digit Span (i.e., recalling a list of numbers in forward and reverse order). However, Bialystok (2011) pointed out that the digit span does not test executive functioning and that the Sun-Moon Stroop task was too easy for the age group being studied (9-12 year olds), and therefore, did not recruit executive functions. Furthermore, Schellenberg's (2011) five EF tasks were uncorrelated, which is surprising given other evidence for relationships among EFs (e.g., Miyake et al., 2000). This suggests that, taken together, these tasks were not suited for measuring overall executive function.

Another study has found that EF does mediate the relation between music lessons and intelligence, and that months of music lessons in children correlate with performance on tasks tapping **inhibition** and **shifting** (Degé et al., 2011). Tasks from the executive function portion of the NEPSY II, a neuropsychological assessment for children, were administered, which included tasks of inhibition (i.e., inhibit saying the shape or direction of an object or arrow), selective attention (i.e., pressing a button whenever hearing the word "red"), planning (i.e., drawing clocks indicating given times), set shifting (i.e., a card sorting task sorting animals by category), and fluency (i.e., drawing as many different design as possible by connecting five dots) with children from the same age group (9-12 year olds) as Schellenberg (2011). Degé and colleagues (2011) found that all of their executive function tasks correlated with the duration of music lessons, and that inhibition and selective attention mediated the relationship between music training and IQ. Given that these two studies (Schellenberg, 2011; Degé et al., 2011) use different categorizations of musicians and non-musicians as well as different EF tasks, it remains uncertain whether there is a direct relationship between musical training and IQ or if it is mediated by improved performance of EFs.

Although some studies did not find relationships between music training and various EF components, there is still evidence supporting the relationship of musical training across all three components of EF. However, generalizations about the link between musical training and EFs cannot be made since these studies use different tasks to measure EF. Furthermore, given the correlational nature of these studies, another limitation to their generalizability is how they define a "musician" versus a "non-musician." For example, Bialystok and DePape (2009) categorized a musician as someone who has studied music for at least half of his or her life and a non-musician as someone who does not have this specialized experience, which contrasts with Schellenberg (2011) who categorized a musician as someone who has taken at least two years of private lessons and a non-musician as someone with fewer than two years of private lessons (see Table 1 for a summary). These varying definitions between studies further preclude the ability to glean generalizable results.

Experimental Evidence

One solution to the issue of how to define a musician is to randomly assign participants to receive music lessons or to a control group. Thus, there has also been a handful of randomized control trials done looking at the effects of music lessons on executive functioning. Results are summarized in Table 2 (adapted from Okada & Slevc, in press).

Bugos, Perlstein, McCrae, Brophy, & Bedenbaugh (2007) randomly assigned older adults to six months of individualized piano lessons or to a no treatment control

Study	Bugos et al., 2007	Moreno et al., 2011	Mehr, 2013 exp 1	Mehr, 2013; exp 2	Roden, 2014	Schellenberg, 2004
Notes on Main / Ancillary Effects	Musicians showed improved performance on Trail Making Test Part B / Digit Symbol	Musicians had larger peak P2 amplitudes on go/no-go task than control group	No overall results when analyzing all tasks, but when looking at 2 spatial tasks, musicians were better at map use/navigation test	No significant differences	Musicians were better at central executive measures than control group	Music group had larger increases in full-scale IQ than control
EF Effect? Effect Size?	**Yes** (d = 0.38)	**Yes** (partial η^2 = .21)	**Yes** (d=.65)	No	**Yes** (Counting Span, partial η^2 =.10; Complex Span, partial η^2 =.12)	**Yes** (d=0.35)
EF Component Tested	Shifting	Inhibition	WM Updating	WM Updating	WM Updating	Reasoning / Problem-Solving
Measures Assessed	Trail Making Test - Part B	ERP - P2 on Go/no-go task	Visual form analysis Map Use/Navigation	Visual Form Analysis Map Use/Navigation	Counting Span Complex Span	WISC-III: Full-scale IQ
Additional Measures	Digit Symbol	Vocabulary & Block Design Subtests from WPPSI-III	PPVT-III Numerical Discrimination	PPVT-III Numerical Discrimination	One-syllable Word Span Non-word Recall Test Corsi Block Test Matrix Span Counting Span Color Span Backwards	K-TEA BASC
Training Length (total hours/duration)	13 hours / 6 months	40 hours / 4 weeks	4.5 hours / 6 weeks	4.5 hours / 6 weeks	58.5 hours / 1.5 years	28 hours / 36 weeks
Music Curriculum	Weekly Individualized Piano Instruction (Theory, Etudes, Lesson); "broad-based music education program"	Computerized music program: primarily listening activities	"Parent-child play" modeled after Eastman Community Music School's Early Childhood Music Program	"Parent-child play" modeled after Eastman Community Music School's Early Childhood Music Program	Weekly lessons on musical instrument of their choice; different instruction based on age	Either standard keyboard lessons or Kodály voice lessons
Control Group	No treament control	Computerized visual arts training: development of visuospatial skills	Visual arts training: "parent-child play" in the style of Reggio Emilia	No treatment control	Natural science training program	Either drama lessons or no lessons
Age of Subjects (SD) Music Group / Control Group	69.6 (4.7) years / 71.4 (6.4)	63.8 months / 63.7 months 5.32 years / 5.31 years 4-6 years	4.86 (.307) years / 4.64 (.268)	4.71 (.260) years / 4.72 (.353)	7.36 (0.57) years / 7.72 (0.68)	Keyboard: 6.20 (.21) years Voice: 6.28 (.21) / Drama: 6.20 (.23) No Lessons: 6.31 (.22)
Initial / Final Sample Size	39 / 31	71 / 48	32 / 29	46 / 45	50	144 / 132

Table 2. Summary of experimental evidence of Musical Training and EFs, adapted from Okada & Slevc (in press).

group, and found that those who received music training had improved performance on the Trail Making Test, Part B, where subjects were required to connect dots alternating between sequential numbers and letters in alphabetical order (i.e., A, 1, B, 2, C, 3, etc.). Although these results need to be interpreted with caution since performance was compared to a no treatment control (Shadish, Cook, & Campbell, 2002; cf. Boot, Simons, Stothart, & Stutts, 2013), these data still suggest that music lessons may lead to benefits in the EF component of **shifting**.

Another study randomly assigned children to a computerized music training program, in which basic music concepts like pitch and rhythm were taught, or a computerized visual arts training program, in which basic concepts like color and shape were taught (Moreno et al., 2011). Although these children did not learn to play an instrument during their training, those who received computerized lessons about music showed better performance than those receiving visual arts training in **inhibitory control**, shown in larger P2 peak amplitudes in no-go trials in a go/no-go task (see Moreno & Farzan, 2015, for discussion).

Two other studies have investigated the effects of music training on working memory **updating**, although they have used different tasks. Roden, Grube, Bongard, and Kreutz (2014) randomly assigned children to 18 months of music lessons on an instrument of their choosing or a natural science training program, and found that those who took music lessons showed better performance on a counting span test and a complex span test. Mehr, Schachner, Katz, and Spelke (2013) randomly assigned children to 6 weeks of either music or visual arts training, and found that who had music lessons showed better performance on a map use/navigation task, which involves holding a 2-D map with landmarks (i.e., circles that represented barrels from an aerial view) in working memory and manipulating it to navigate a 3-D world (i.e., real barrels in the room) (Mehr et al., Exp 1). However, note that Mehr et al. (2013) did not find differences on an omnibus test of all four tasks assessed between the music and visual arts group, but found this difference only when analyzing data from the two spatial tasks described here. Moreover, in a follow-up study comparing a new group of children randomly assigned to either music lessons or to a no treatment control, this effect was not found (Mehr et al., 2013, Exp 2).

In sum, there appears to be an effect of musical training across all three components of EF, however it is hard to generalize these results since the type and extent of musical training given to participants was varied. Although these experimental studies can provide better tests of the causal relationship between musical training and EFs, correlational studies are still an important first step in identifying what links musical training has with EFs. Significant correlational evidence may be evidence for transfer effects and can inform what exactly should be targeted in future experimental training studies, or it could reflect exaggerated pre-existing individual differences for those who choose to take music lessons.

Study Objectives

The goal of this study is to more thoroughly investigate whether musical training is associated with EFs by attempting to improve upon some of the aforementioned shortcomings in previous studies. To improve upon the "task impurity problem," this study uses multiple tasks to assess each EF component (Miyake et al., 2000). By using a latent variable approach, I can estimate what is common between tasks of each EF component, and get a better estimate of the underlying component of interest removed from task-specific effects. Additionally, this battery of EF tasks has previously been validated and used in other research looking at individual differences in EF abilities (Friedman et al., 2006; Friedman et al., 2008; Friedman, Miyake, Robinson, & Hewitt, 2011; Miyake et al., 2000; Ito et al., 2015). To improve upon the problem of binary categorization of “musicians” and “non-musicians,” this study uses a continuous measure of musical training from a well-validated questionnaire (described in detail in the Musical Training Measures section below), in lieu of arbitrarily categorizing participants by number of years of music lessons.

Method

Participants

161 subjects (total, *N* = 150 after list-wise deletion of missing data) were recruited from the University of Maryland undergraduate research pool. A target sample size of 150-200 was set based on similar participant numbers in other studies examining individual differences in cognitive processes (137 in Miyake et al., (2000); 133 in Engle, Tuholski, Laughlin & Conway (1999); 120 in Conway, Cowan, Bunting, Therriault, & Minkoff (2002); 215 in Shipstead, Lindsey, Marshall, & Engle, (2014)), while still being feasible to collect over two semesters. Participants reported normal hearing, were not colorblind, and were native English speakers. Participants needed to be able to see color in order to complete the Stroop task (described in detail below), in which participants saw colored words and named the color of the font. They also needed to be native English speakers in order to control for the different Stroop effects found in unbalanced bilinguals when speaking in their dominant versus non-dominant language (Rosselli et al., 2002). Participants either completed testing for course credit or monetary compensation.

Measures

EF Measures

Participants completed a battery of nine EF tasks measuring inhibition, switching, and updating abilities (see Table 3 for a summary). This battery of tasks has previously been validated and used in other research looking at individual differences in EF abilities (Friedman et al., 2006; Friedman et al., 2008; Friedman, Miyake, Robinson, & Hewitt, 2011; Ito et al., 2015; Miyake et al., 2000).

Inhibition	**Updating**	**Switching**
Stop-signal*:* See square or circle and indicate shape, but inhibit this response when a beep sounds (on 25% of trials).	**Keep track:** See words presented serially and remember the most recently presented word belonging to each of 2-5 target categories.	**Number-Letter:** Categorize letter-number pairs by letter (vowel/consonant) when on one side and by number (odd/even) when on the other
Antisaccade*:* See a cue, then respond to a briefly presented target that appears on the opposite side of screen.	**Letter Memory:** See letters presented serially (with variable list lengths) and maintain the last four letters.	**Color-Shape:** Categorize colored shapes by color (red/green) or shape (circle/triangle), as indicated by a cue.
Stroop: Name the font color of letter strings that are incongruent color words (e.g., “blue” in red font) or are strings of asterisks.	**Spatial N-back:** See a series of spatial locations indicated serially and indicate if the location is the same location indicated n series earlier.	**Category-Switch:** Categorize words as living or non-living, or as larger or smaller than a soccer ball, as indicated by a cue.

Table 3. Summary of EF task battery.

Inhibition Tasks. These three tasks all require the inhibition of different prepotent responses despite otherwise distinct task demands.

Antisaccade. In this task adapted from Roberts, Hager, & Heron (1994), participants first saw a fixation cross in the center of the screen, then saw a cue to either the left of the right side of the fixation cross. After this cue, a numeric target (a number 1 – 9) appeared for 150 ms, and the participant was asked to verbally indicate the target number. In the first prosaccade block, the cue and target appeared on the same side of the screen in order to create a prepotent response to this stimulus. Next, participants completed three antisaccade blocks, in which the target appeared on the opposite side of the screen of the cue. Performance was measured by the proportion of correct responses in the antisaccade blocks.

Stop Signal. This task from Verbruggen, Logan, & Stevens (2006) consisted of participants fixating on a cross in the center of the screen, which was replaced by either a square or a circle. They were instructed to push left as quickly as possible if they saw a square, and were instructed to push right as quickly as possible if they saw a circle. On a quarter of the trials, participants saw the shape then heard a stop signal (a beep from the computer), and were instructed to withhold any response upon hearing a stop signal. On each trial in which the participant received a stop signal, the onset of the stop signal was adjusted until participants could correctly inhibit 50% of the responses. Performance was measured as the stop signal reaction time (SSRT), an estimate of how long it would take to inhibit an “already-initiated response” (Logan, 1994). This was calculated by finding the difference between the median reaction time for identifying the shape and the average onset time of the stop signal. This provided a measure of how much time each subject needed to accurately inhibit responding.

Stroop. In this task, participants read color words (i.e., red, blue, or green) presented on a screen, and were instructed to say the color of the font aloud (Stroop, 1935). There were congruent trials, in which the color of the font was the same as the word (e.g., the word “blue” written in blue ink), incongruent trials, in which the color of the font did not match the word (e.g., the word “blue” written in red ink), as well as neutral trials, in which a string of asterisks appeared (e.g., “****” written in blue ink). The Stroop effect was the difference in means between incongruent trials and neutral trials for correct responses.

Updating Tasks. These three tasks all require that items in working memory are constantly being added or deleted given different task demands.

Keep track. In this task, adapted from Yntema (1963), participants kept track of exemplars from six different categories (i.e., relatives, countries, colors, animals, metals, and distances). In each trial, participants were presented with 2 to 5 categories (for a total of 4 different difficulty levels), then 15-25 words belonging to those categories were shown one at a time. Participants were instructed to verbally recall the most recent exemplar they saw from each of the categories presented in the trial. Performance was calculated as the proportion of correct exemplars recalled.

Letter memory. In this task, adapted from Morris and Jones (1990), participants saw a string of letters (consonants only) appear on the screen one at a time. The strings were either 9, 11, or 13 letters long. For all trials, participants were tasked with saying aloud only the last four letters in the string after each letter appeared, thereby constantly updating which letters were being held in memory. Performance was calculated as the proportion of accurate strings said aloud.

Spatial n-back. In this task (from Friedman et al., 2008), one of twelve, stationary boxes on a screen flashed black, and participants were instructed to indicate whether or not that same box had flashed previously. Participants completed a 2-back condition, in which they indicated whether or not the same box flashed 2 trials earlier as well as a 3-back condition, in which they indicated whether or not the same box had flashed 3 trials earlier. Performance was calculated as the proportion of correct responses across both conditions.

Shifting Tasks. These three tasks all require set shifting between two types of binary categorization with distinct task demands.

Number-Letter. In this task (adapted from Rogers & Monsell, 1995), participants saw a letter-number pair appear in one of four quadrants on the screen. If the pair appeared in one of the two top quadrants, the participant was instructed to categorize the number as odd or even, and if the pair appeared in one of the two bottom quadrants, the participant was instructed to categorize the letter as a consonant or vowel. Performance was calculated as the switch cost, which is the difference in mean reaction time between switch trials (where participants switched between what they were categorizing) and repeat trials (where participants categorized numbers or letters twice in a row).

Color-Shape. In this task (from Miyake, Emerson, Padilla, & Ahn, 2004), participants first saw a cue (“C” for color or “S” for shape), then saw red or green circles and triangles. If the participant saw a “C,” they needed to indicate if the color was red or green, and if they saw an “S,” they needed to indicate if the shape was a circle or a triangle. Performance was calculated as the switch cost, the difference in reaction time between switch trials and repeat trials.

Category switch. In this task, (adapted from Mayr & Kliegl, 2000), participants were again asked to categorize stimuli by one of two dimensions. Examples of stimuli include “alligator,” “coat,” “lion,” and “knob.” Participants first saw a cue (i.e., a heart or crossed arrows), then a stimulus. If they saw a heart, they needed to indicate if the stimulus was living or non-living, and if they saw crossed arrows, they needed to indicate if the stimulus was larger or smaller than a soccer ball. Performance was calculated as the switch cost.

Musical Training Measures

Gold-MSI. Musical training was assessed with the Goldsmith Musical Sophistication Index (Gold-MSI) (Müllensiefen, Gingras, Musil, & Stewart, 2014). The Gold-MSI is a self-report questionnaire that measures general “musical sophistication” with questions in five subscales: active engagement, perceptual abilities, musical training, singing abilities, and emotions. This inventory also has shown high internal consistency and has been validated through comparisons with a standard musical ability discrimination test and another musical self-report inventory (i.e., the Musical Engagement Questionnaire, Werner, Swope, & Heide, 2006). Measures were collected from all subscales, however only the musical training subscale was used in this analysis. The musical training subscale differs from previous ways used to measure musical experience in that it contains seven questions regarding musical training, which include: years of instrument training, years of music theory training, regular daily practice, the number of hours practiced at peak of interest, the number of instruments played, whether compliments about music performances have been received, and whether he/she considers himself/herself a musician. Since this measure takes into account how long one has taken music lessons as well as the intensity of practice, participants' scores from the musical training subtest provided a continuous and more robust measure of musical training (rather than only looking at duration of music lessons like most studies), which was used to predict performance on the EF tasks.

Covariate Measures

Measures of socioeconomic status (SES), handedness, and intelligence (IQ) were collected in order to control for other variables that are correlated with music training and with EFs.

SES. Because musical participation is unevenly distributed across SES (Southgate & Roscigno, 2009) and SES is a predictor of EF ability (Hackman & Farah, 2009), the MacArthur Scale of Subjective Social Status (Adler & Stewart,

2007) was administered. Here, participants indicated where they believed they stood (in terms of money, education, and job status) relative to others in the U.S., on a scale of 1 to 10.

Handedness. Because handedness has been associated with performance in EF tasks (Beratis, Rabavilas, Kyprianou, Papadimitriou, & Papageorgiou, 2013) and aspects of musical ability (Kopiez, Galley, & Lee, 2006), the Edinburgh handedness inventory (Oldfield, 1971) was administered. This questionnaire asks which hand one prefers to use when doing various activities (e.g., writing, drawing, or using scissors), and provides a continuous measure of laterality, scored on a scale from -100 (completely left handed) to +100 (completely right handed).

IQ. Since IQ is linked with both EF ability (Friedman et al., 2006) and music training (Schellenberg, 2006), the Cattell Culture Fair Intelligence Test (Scale 3B) was administered to measure fluid intelligence (Cattell & Cattell, 1960). This test contained four subtests, in which subjects completed a sequence of drawings, classified which images were different from others, completed a matrix of patterns, and chose which option of geometric drawings satisfied a given rule. Scores were calculated as the proportion of correct answers.

Procedure

Participants completed two sessions of 1.5 hours each, which were separated by at least a day. The order of the tasks was fixed as follows: stop signal, spatial 2-back, category switch, Stroop, keep track, color-shape, letter memory, antisaccade, number-letter, and spatial 3-back (following Ito et al., 2015). In this way, no sequential tasks tapped the same EF component. All EF tasks were programmed and presented in Psyscope (Cohen, MacWhinney, Flatt, & Provist, 1993) or Tscope and C (Verbruggen, Logan, & Stevens, 2008). After the first five EF tasks were completed during the first session, participants filled out questionnaires (Gold-MSI, Müllensiefen et al., 2014; SES, Adler & Stewart, 2007; Handedness, Oldfield, 1971). And after all of the EF tasks were completed in the second session, participants completed the Cattell Culture Fair Intelligence Test (Cattell & Cattell, 1960).

Statistical Procedures

All scores were standardized (z-scored). To facilitate interpretation, scores for each task were adjusted so that larger scores meant better performance (i.e., z-scaled scores for all three shifting tasks, stop signal task, and the Stroop task were multiplied by negative 1).

Data Trimming. For all reaction time (RT) measures, data on accurate trials were trimmed to exclude all RTs under 200 ms and above 3000 ms. Then, RTs that were 2 standard deviations away from each participant’s mean were excluded. For all three shifting tasks, trials following an incorrect response were dropped because a previous error reflected that a switch in categorization was not achieved (following Friedman et al., 2008).

Results

Descriptive statistics for participant demographics, covariate measures, and the musical training measure as well as descriptive statistics for each of the executive function tasks are provided in Tables 4 and 5.

Measure	Mean	SD	Min	Max	Skewness	Kurtosis
Age	19.27	1.11	17	22	0.48	-0.6
Cattell IQ	27.92	4.64	11	38	-0.43	0.34
SES	6.45	1.52	3	9	-0.28	-0.58
Handedness	68.09	43.19	-100	100	-2.5	6.15
Musical Training	25.38	10.2	7	47	0.08	-0.91

Table 4. Descriptive statistics for participant demographics, covariate measures, and musical training measure.
N = 152 for all measures except SES (*N* = 150) and Age (*N* = 147)

Task	Mean	SD	Min	Max	Skewness	Kurtosis
Antisaccade[a]	0.73	0.14	0.34	0.97	-0.6	-0.01
Stop Signal[b]	269.04	35.72	177.3	420.6	0.64	2.02
Stroop[c]	119.88	61.93	-37.18	320.34	0.713	0.78
Keep Track[a]	0.73	0.09	0.45	0.91	-0.48	0.21
Letter Memory[a]	0.77	0.14	0.38	1	-0.26	-0.37
Spatial N-Back[a]	0.8	0.07	0.53	0.96	-0.5	1.11
Number-Letter[d]	176.83	113.55	-39.15	628.29	1.13	2.12
Color-Shape[d]	151.62	135.82	-69.02	760.37	1.79	4.31
Category Switch[d]	118.89	92.59	-39.49	432.45	0.88	0.6

Table 5. Descriptive statistics for each EF task.

Total *N* = 152 (including two participants who did not respond to the SES measure)
[a] Proportion Accuracy
[b] milliseconds, measured as the stop-signal reaction time
[c] milliseconds, measured as incongruent trials minus neutral trials
[d] milliseconds, measured as switch trials minus stay trials

Model Estimation

The first step in analysis was to conduct a confirmatory factor analysis (CFA) to ensure the nine executive function tasks fit well onto the three-factor model of EFs as seen in previous studies using the same task battery (Miyake et al., 2000; Friedman et al., 2006; 2008; Ito et al., 2015). Our model contains three latent factors of EF (i.e., inhibition, shifting, and updating), which are each measured by three tasks. Conducting CFA allows us to assess factor loadings for each task, which are values that tell us how much the underlying latent factor influences each task score. In Figure 1, standardized factor loadings are shown on one-headed arrows pointing from latent factor (circles) to each measured task (squares). CFA also provides overall model fit estimates that tell us if the data fit our hypothesized three-factor model.

A three-factor model was estimated using Mplus version 7.4 (Muthen & Muthen, 2015). First, we look at the Chi-Square Test of Model Fit (χ^2), the likelihood ratio, which tests the null hypothesis that the observed data are no different from the expected population covariance matrix from the model (i.e., that the model fits the data). The alternative hypothesis is that our observed data do not fit the population covariance specified by our model. Thus, a non-significant χ^2 means that the model fits the data well. This model fit index is supplemented by three other model fit indices: the Comparative Fix Index (CFI), Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). The CFI is classified as a comparative index of model fit because it indicates "improvement in model fit by comparing the hypothesized model in which structure is imposed with the less restricted nested baseline model" (Byrne, 2013, p. 72). A CFI above .95 is considered good model fit. The RMSEA and SRMR are classified as absolute indices of model fit because they do not compare the hypothesized model with a "reference model in determining the extent of model improvement; rather, they depend only on determining how well the hypothesized model fits the sample data" (Byrne, 2013, p. 72). A RMSEA less than .05 and a SRMR under .08 show good model fit. All nine tasks loaded significantly onto their EF constructs, and there was good model fit (χ^2(24, N=152) = 31.18, *p* = .15; *CFI* = .96; *RMSEA* = .04; *SRMR* = .05).

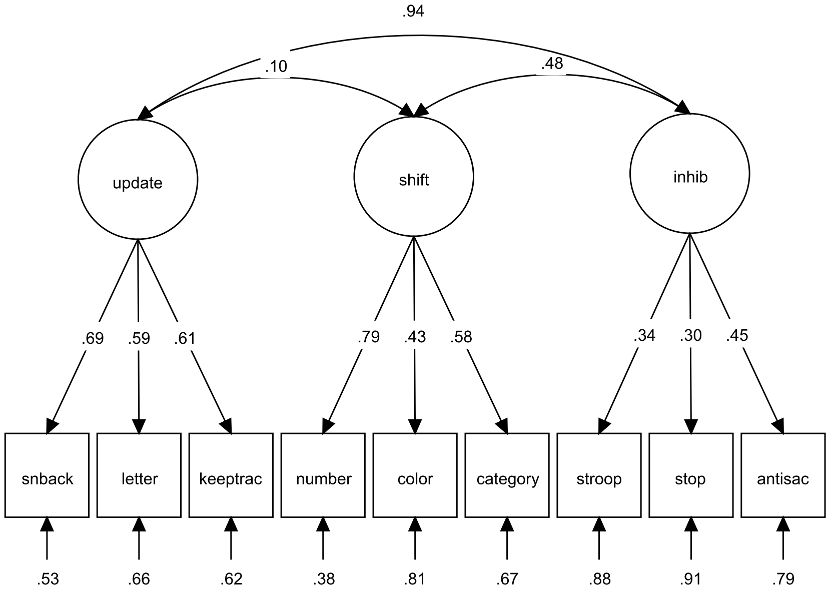

Figure 1. Three-factor model of EFs. Unobserved latent factors are in circles and measured tasks are in squares. The single-headed arrows from latent factors to measured tasks are standardized factor loadings, which are all significant (p<.05). The short arrows on the bottom represent measurement error – squaring this number gives the amount of variance in each task left unaccounted for by latent factor. The double-headed arrows on the top represent correlations between latent factors.

Linear Mixed-Effects Modeling

Since it is not feasible to run an individual differences analysis in a SEM model given the limited sample size, linear mixed-effects models were used. Linear mixed effects models do not require such large sample sizes, but can still estimate effects for each latent construct (following von Bastien & Oberauer, 2013). Additionally, linear mixed effects models allow for the specification of fixed and random effects, which control for systematic sources of variance. Fixed effects model effects of independent variables or experimental conditions, and random effects estimate individual differences due to random sampling from the population.

To capture effects in a way analogous to latent measures in SEM, we modeled task as a random effect nested within *EF type* (i.e., inhibition, shifting, and updating). This essentially treats tasks as a sample drawn from a population of tasks measuring each construct of interest. The model also included *subjects* as a random effect, and the covariate measures and musical training as fixed effects. Analyses were run in R (v. 3.2.4; R Core Team, 2016) with the package "lme4" (v. 1.11.1; Bates, Maechler, Bolker, & Walker, 2015). Since degrees of freedom are not estimated in mixed effects models (and there are various ways to estimate p-values), a t-value more extreme than ± 2 will be determined significant (Gelman & Hill, 2007).

In order to assess whether musical training could predict performance differentially for EF components, a linear mixed effects model was run with *subjects* and *tasks* as random effects, and *EF type*, *music training*, and the covariate measures[1] as fixed effects. The random effects structure included random intercepts for *subjects* and by-subject random slopes for *EF type*. Although *task* was entered as a random effect, its random intercept had to be removed for the models to converge (presumably because all variables were z-scored and so no intercepts differed from

---

[1] Handedness and SES did not significantly improve model fit, but because they were predicted a priori to be associated with musical training and planned as covariates, they are nonetheless included in the models reported below.

zero); however the models did include by-task random slopes for *Musical Training*. *EF type* was contrast coded so that the first contrast compared shifting scores with the mean of inhibition and updating scores, and the second contrast compared inhibition and updating scores only.

As reported in Table 6, IQ was a significant predictor of task performance, and there were significant interactions of music training by EF type, which both indicate that the extent to which musical training predicts task scores differs as a function of EF type. Table 6 summarizes the results.

Table 6. Mixed effects model examining the effects of musical training and EF type on EF performance.

Parameters	Fixed effects				Random effects			
					By Subjects			By Tasks
	Estimate	*SE*	*t*		*SD*	EF1	EF2	*SD*
Intercept	-0.01	0.04	-0.15		0.36	-0.19	0.09	--
Musical Training	0.03	0.04	0.72		0.01	--	--	0.01
EFType1	0.01	0.07	0.144		0.61	--	-0.79	--
EFType2	<.01	0.03	<.01		0.13	--	--	--
IQ	0.21	0.04	5.1	*	--	--	--	--
Handedness	-0.04	0.04	-1.04		--	--	--	--
SES	0.01	0.04	0.23		--	--	--	--
MusicalTraining*EFType1	0.3	0.07	4.29	*	--	--	--	--
MusicalTraining*EFType2	-0.06	0.03	-2.12	*	--	--	--	--

Note. Factors were contrast coded as follows: EFType1 (-1 = Shifting, .5 = Inhibition, .5 = Updating), EFType2 (0 = Shifting, .5 = Inhibition, -.5 Updating). Under Random effects, values to the right of the SD columns indicate estimated correlations between random effects. Model formula for correlated random effects model: TaskPerformance ~ MusicalTraining*EFtype + IQ + Handedness + SES + (0+MusicalTraining|Task) + (EFtype|Subject)
* |*t*| > 2.0, indicating a significant effect (Gelman & Hill, 2007).

To unpack the interactions found in the omnibus model, separate linear mixed-effects models were conducted for each of the three EF components. For each model, the same structure of mixed-effects models was used except that random slopes for EF type for Subject random effect were taken out (since only one EF type was present in each model). These separate models suggest that the interaction effects above reflect that musical training significantly predicts updating ability.

LME – Inhibition

Looking at only the three inhibition task scores, musical training did not predict a significant amount of variance in scores (estimate = .076, *t* = 1.08), however IQ was a significant predictor (estimate = .17, *t* = 3.18).

Table 7. Mixed effects model for inhibition task scores.

Parameters	Fixed effects				Random effects	
					By Subjects	By Tasks
	Estimate	*SE*	*t*		*SD*	*SD*
Intercept	-0.006	0.05	-0.13		0.28	--
Musical Training	0.08	0.07	1.08		--	0.08
IQ	0.17	0.05	3.18	*	--	--
Handedness	-0.05	0.05	-0.91		--	--
SES	0.03	0.05	0.54		--	--

Note. Model formula for correlated random effects model:
InhibitionTaskPerformance ~ MusicalTraining + IQ + Handedness + SES + (0+MusicalTraining|Task) + (1|Subject)
* |*t*| > 2.0, indicating a significant effect (Gelman & Hill, 2007)

LME – Shifting

Looking at only the three shifting task scores, musical training did not predict a significant amount of variance in scores (estimate = -.10, *t* = -1.66). Furthermore, none of the covariates predicted a significant amount of variance.

Table 8. Mixed effects model for shifting task scores.

Parameters	Fixed effects				Random effects	
					By Subjects	By Tasks
	Estimate	*SE*	*t*		*SD*	*SD*
Intercept	-0.01	0.06	-0.24		0.59	--
Musical Training	-0.11	0.07	-1.66		--	0.00
IQ	0.03	0.07	0.44		--	--
Handedness	0.01	0.07	0.44		--	--
SES	-0.04	0.06	-0.69		--	--

Note. Model formula for correlated random effects model: ShiftingTaskPerformance ~ MusicalTraining + IQ + Handedness + SES + (0+MusicalTraining|Task) + (1|Subject)

LME – Updating

Looking at only the three updating task scores, musical training accounted for a significant amount of variance in scores (estimate = .13, *t* = 2.40) as well as IQ (estimate = .37, *t* = 6.59).

Table 9. Mixed effects model for updating task scores.

Parameters	Fixed effects				Random effects	
					By Subjects	By Tasks
	Estimate	*SE*	*t*		*SD*	*SD*
Intercept	0.004	0.05	0.07		0.21	--
Musical Training	0.13	0.06	2.40	*	--	0.00
IQ	0.37	0.06	6.59	*	--	--
Handedness	-0.07	0.05	-1.34		--	--
SES	0.03	0.05	0.55		--	--

Note. Model formula for correlated random effects model: UpdatingTaskPerformance ~ MusicalTraining + IQ + Handedness + SES + (0+MusicalTraining|Task) + (1|Subject)
* |*t*| > 2.0, indicating a significant effect (Gelman & Hill, 2007).

Discussion

Overall, these data show a positive relationship between musical training and working memory updating, but no relationship between musical training and inhibition or shifting. These results suggest that the significant relationship between musical training and shifting found in previous studies may reflect a failure to control for other confounding variables (e.g., Hanna-Pladdy & McKay, 2011; Zuk et al., 2014) or possibly a function of a binary versus continuous measure of musical training.

Similarly, previous studies finding support for the relationship between musical training and inhibitory control did not use IQ as a covariate when assessing the relationship with inhibition tasks (Bialystok & DePape, 2009; Degé et al., 2011; Moreno et al., 2011); however, some studies did measure IQ and found it to be the same across groups and therefore did not use it in subsequent analyses (Bialystok & DePape, 2009; Moreno et al., 2011). To see if a positive association between musical training and inhibition ability existed here, a model with only a fixed effect of musical training (i.e., with no covariates) was run, and musical training was almost able to predict a significant amount of variance in inhibition tasks (estimate = .14, t = 1.99; although the full model with covariates had a better fit ($\chi^2(3) = 11.90$, p = .008)). Musical training and IQ were also found to correlate with performance on multiple tasks (see Appendix B for a summary), which further highlights the need for future studies to control for extraneous variables as well. Fitting with results from Bialystok and DePape (2009), there was a significant zero-order correlation between music training and performance on the Stroop task (b = -1.43, SE = 0.48, t = -2.95, p = 0.004); however, this relationship disappeared when controlling for covariates (b = -0.96, SE = 0.52, t = -1.86, p = 0.065). In contrast with Moreno et al., (2011), the current study surprisingly did not find a significant zero-order correlation between musical training and performance on the Stop Signal task, which is similar to the go/no-go task used in Moreno et al. (2011). This may have been due to different outcome measures used: this study used stop-signal reaction time, while Moreno et al. (2011) found an effect with task accuracy.

Musical training was able to predict updating scores even after accounting for the (significant) amount of variance predicted by IQ. This fits well with previous evidence that musicians outperform non-musicians on N-back tasks (Oechslin et al., 2013; Pallesen et al., 2010; Slevc et al., 2016). One reason for this relationship might be the demands of reading music, especially sight-reading. Sight-reading music, or playing unpracticed from a score, requires looking ahead in the music to prepare for what will be played. Good sight-readers typically look about four notes ahead of where they are playing (Furneaux & Land, 1999; cf. Drake & Palmer, 2000; Goolsby, 1994). Furthermore, sight-reading ability has also been correlated with eye-hand span (i.e., the number of notes played after sheet music is taken away) (Sloboda, 1974). This suggests that expert sight-readers look farther ahead in the music and constantly update the contents of working memory: they must keep in mind which notes are being played and which are yet-to-be played. Correspondingly, working memory capacity (as indexed by an average of scores for operation span, reading span, rotation span, and matrix span) is associated with sight-reading ability (Meinz & Hambrick, 2010).

In fact, some exploratory analyses found that for participants who reported that they could read music and self-reported their level of sight-reading ability[2], musical training better predicted updating scores for those who were better sight readers (i.e., a significant interaction between musical training and self-rated sight reading ability on updating scores: estimate = .25, $t = 2.41$, $N = 143$). Furthermore, looking only at the subset of participants who reported they could *not* read music ($N = 60$), musical training no longer significantly predicted performance on the updating tasks. This suggests that reading and sight-reading may indeed play some role in musical training's relationship with EF (cf. Meinz & Hambrick, 2010). However, these results are from exploratory analyses and it is unknown how accurately participants can self-report sight-reading ability, so future studies investigating this link should measure this construct more thoroughly.

Surprisingly, we did not find a significant relationship between musical training and SES. Our subjective measure of SES was significantly correlated with objective measures of SES (i.e., with both parental education and parental income[3]), which fits with previous findings that subjective and objective measures of SES are related, but not perfectly correlated (Adler & Stewart, 2007). So perhaps this non-relationship may reflect limited variability within our sample.

---

[2] These exploratory analyses were based on additional questionnaire items: 1) Can you read music? (Yes or No). 2) If yes, how well are you able to sight-read music (playing music you are seeing for the first time)? (Scale from 1-5)

[3] Measures of parental income and education for mothers and fathers were collected on a 9-point scale.

General Discussion

There is a slew of evidence linking musical training to a wide range of cognitive processes, but it has been difficult to assess the strength of this evidence given that most studies use single tasks to measure a single cognitive process with varying definitions of “musicians” and “non-musicians.” This study sought to improve upon these problems by measuring multiple tasks tapping multiple constructs of EF (inhibition, shifting, and updating) as well as using a continuous measure of musical training in order to provide a clearer picture of the relationship between musical training and executive function.

Individual differences in musical training were able to predict updating ability (estimated from scores on Keep Track, Letter Memory, and N-back tasks), but were not predictive of inhibitory control ability (estimated from scores on Stroop, Stop-Signal, and Antisaccade tasks) or shifting ability (estimated from scores on Color-Shape, Category Switch, and Number-Letter tasks). Musical training’s relationship with memory updating falls in line with previous correlational work (Oechslin et al., 2013; Pallesen et al., 2010; Slevc, et al., 2016; Zuk et al., 2014) as well as experimental studies (Mehr et al., 2013, Experiment 1; Roden et al., 2014), and may be attributable to the association between sight-reading ability and working memory (Meinz & Hambrick, 2010). However, since this is a correlational study, the directionality of this relationship is still unclear. One possibility is that music training affects working memory updating abilities, and perhaps updating ability is further improved with the practice of reading and sight-reading music. Another possibility is that those with higher working memory updating abilities choose to go into music lessons. A third possibility is that music lessons might exaggerate these pre-existing differences (Schellenberg, 2015).

Although this study cannot warrant causal claims for musical training, this significant relationship between musical training and working memory updating (and non-significant relationships with inhibition and shifting) could inform work on the transfer effects of musical training by indicating appropriate avenues to explore in experimental studies that (ideally) randomly assign participants to musical training or a well-matched control training regimen. Conversely, it can also inform work on individual differences and pre-existing differences between those who do and do not take music lessons. If different aspects of musical training draw on EFs, this can also inform future work on who is more likely to excel in musical training. In sum, the relationship between musical training and EFs is complex, and the present results help to add clarity to the large body of work investigating music lessons and non-musical cognitive abilities.

Appendices

Appendix A. Correlation matrix of EF tasks.

	Inhibition			Updating			Switching		
	Anti-saccade	Stop Signal	Stroop	Keep Track	Letter Memory	Spatial N-back	Number-Letter	Color-Shape	Category Switch
Antisaccade									
Stop Signal	0.11								
Stroop	0.26**	-0.04							
Keep Track	0.24**	0.24**	0.15						
Letter Memory	0.29***	0.20*	0.28***	0.31***					
Spatial N-Back	0.27***	0.16	0.21*	0.47***	0.39***				
Number-Letter	0.18*	0.15	0.09	0.12	0.002	0.06			
Color-Shape	-0.04	0.05	-0.03	0.04	-0.14	0.03	0.36***		
Category Switch	0.16*	0.12	0.18*	0.01	-0.01	0.1	0.44***	0.25**	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. Scores have been scaled so that larger number indicate better performance (switch costs, stop-signal ssrt, and Stroop effect scores have been multiplied by -1)

Appendix B. Zero-order correlations between musical training, covariates, and EF task scores.

	Musical Training	SES	IQ	Handedness
Musical Training				
SES	-0.04			
IQ	0.33***	0.17*		
Handedness	-0.13	0.06	0.01	
Antisaccade	0.17*	0.09	0.29***	0.02
Category Switch	-0.12	-0.08	0.02	0.1
Color-Shape	-0.07	0	-0.03	-0.13
Keep Track	0.23**	-0.07	0.38***	-0.06
Letter Memory	0.30***	0.13	0.41***	-0.15
Number-Letter	-0.11	-0.02	-0.03	0.1
Spatial N-back	0.27***	0.19*	0.48***	-0.04
Stop Signal	0.01	0.04	0.12	-0.05
Stroop	0.23**	0.03	0.19*	-0.12

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. Scores have been scaled so that larger number indicate better performance (switch costs, stop-signal ssrt, and Stroop effect scores have been multiplied by -1)

References

Adler, N., & Stewart, J. (2007). The MacArthur scale of subjective social status. *MacArthur Research Network on SES & Health. Retrieved from http://www.macses.ucsf. edu/Research/Psychosocial/subjective.php*.

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *22*(2), 366-377.

Bates, D., Maechler, M., Bolker, B., & Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01.

Benz, S., Sellaro, R., Hommel, B., & Colzato, L. S. (2015). Music Makes the World Go Round: The Impact of Musical Training on Non-musical Cognitive Functions—A Review. *Frontiers in Psychology*, *6*, 2023. http://doi.org/10.3389/fpsyg.2015.02023

Beratis, I. N., Rabavilas, A. D., Kyprianou, M., Papadimitriou, G. N., & Papageorgiou, C. (2013). Investigation of the link between higher order cognitive functions and handedness. *Journal of Clinical and Experimental Neuropsychology*, *35*(4), 393-403.

Berlin, B., & Markow, A. (2002). *Four star sight reading and ear tests: Daily exercises for piano students* (Books 6, 8, and 10; S. McBride Smith, Series Ed.). Mississauga, Ontario, Canada: Frederick Harris.

Bialystok, E. (2011). Commentary: How does experience change cognition?

Evaluating the evidence. *British Journal of Psychology, 102*, 303–305.

Bialystok, E., & DePape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565.

Bilhartz, T. D., Bruhn, R. A., & Olson, J. E. (2000). The Effect of Early Music Training on Child Cognitive Development. *Journal of Applied Developmental Psychology*, *20*(4), 615–636.

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The Pervasive Problem With Placebos in Psychology: Why Active Control Groups Are Not Sufficient to Rule Out Placebo Effects. *Perspectives on Psychological Science*, *8*(4), 445–454. http://doi.org/10.1177/1745691613491271

Broader Minded Media Resources. (2015). Retrieved May 1, 2015, from http://www.nafme.org/take-action/advocacy-resources/broader-minded-media-resources/

Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007). Individualized Piano Instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, *11*(4), 464–71. http://doi.org/10.1080/13607860601086504

Bugos, J., & Mostafa, W. (2011). Musical Training Enhances Information Processing Speed. *Bulletin of the Council for Research in Music Education*, *187*, 7–18.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge: New York.

Butzlaff, R. (2000). Can Music Be Used to Teach Reading? *Journal of Aesthetic*

Education, *34*(3), 167–178. http://doi.org/10.2307/3333642

Cattell, R. B., & Cattell, A. K. S. (1960). *Cattell Culture Fair Intelligence Test: A Measure of" g"*. Bobbs-Merrill.

Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature*, *396*(6707), 128. http://doi.org/10.1038/24075

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments.*Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257-271.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*,*30*(2), 163-183.

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369-382.

Deary, I. J. (2001). Intelligence: A very short introduction. Oxford, UK: Oxford University Press.

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: a relation mediated by executive functions. *Music Perception*, *29*(2), 195-201.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135–168.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool Program Improves Cognitive Control. *Science*, *318*(5855), 1387–1388.

http://doi.org/10.1126/science.1151148

Dixon, S. (2002) MidiCompare (computer software). Vienna: Austrian Institute for Artificial Intelligence.

Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2015). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review, 22*(3), 1-11.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, *74*(1), 1–32. http://doi.org/10.1016/S0010-0277(99)00061-X

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172-179.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201-225.

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: a behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410-1430.

Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1436), 2435–2440. http://doi.org/10.1098/rspb.1999.0943

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. New York: Cambridge University Press.

Goolsby, T. W. (1994). Eye Movement in Music Reading: Effects of Reading Ability, Notational Complexity, and Encounters. *Music Perception*, *12*(1), 77–96.

Hamers, J. F., & Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior*, *11*(3), 303-310.

Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, *25*(3), 378.

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

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music Training Improves Verbal but Not Visual Memory: Cross-Sectional and Longitudinal Explorations in Children. *Neuropsychology*, *17*(3), 439–450. http://doi.org/10.1037/0894-4105.17.3.439

Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, *108*(2), 187.

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*, *105*(19), 6829-6833.

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time Tagging: A Key to Musicians' Superior Memory. *Music Perception*, *20*(3), 307–313. http://doi.org/10.1525/mp.2003.20.3.307

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*(1), 41–56.

Kirschner, S. & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. http://doi.org/10.1016/j.evolhumbehav.2010.04.004

Kopiez, R., Galley, N., & Lee, J. I. (2006). The advantage of a decreasing right-hand superiority: The influence of laterality on a selected musical skill (sight reading achievement). *Neuropsychologia*, *44*(7), 1079-1087.

Kopiez, R., & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*(01), 97-120.

Lamb, S. J., & Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Educational Psychology*, *13*(1), 19–27.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59-80.

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *The Quarterly Journal of Experimental Psychology*, *64*(11), 2153–2167. http://doi.org/10.1080/17470218.2011.603427

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring Individual and Joint Action Outcomes in Duet Music Performance. *Journal of Cognitive Neuroscience*, *25*(7), 1049–1061.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language*. San Diego: Academic Press.

Logue, S. F., & Gould, T. J. (2014). The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition. *Pharmacology Biochemistry and Behavior*, *123*, 45-54.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1124–1140.

Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two Randomized Trials Provide No Consistent Evidence for Nonmusical Cognitive Benefits of Brief Preschool Music Enrichment. *PLoS ONE*, *8*(12), 1–12. http://doi.org/10.1371/journal.pone.0082007

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate Practice Is Necessary but Not Sufficient to Explain Individual Differences in Piano Sight-Reading Skill: The Role of Working Memory Capacity. *Psychological Science*, *21*(7), 914–919. http://doi.org/10.1177/0956797610373933

Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270-291.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167-202.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, *115*(2), 123-142.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. http://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Moore, G. P., & Chen, J. (2010). Timings and interactions of skilled musicians. *Biological Cybernetics*, *103*(5), 401–414. http://doi.org/10.1007/s00422-010-0407-5

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-Term Music Training Enhances Verbal Intelligence and Executive Function. *Psychological Science*, *22*(11), 1425–33. http://doi.org/10.1177/0956797611416999

Moreno, S., & Farzan, F. (2015). Music training and inhibitory control: a multidimensional model. *Annals of the New York Academy of Sciences*,*1337*(1), 147-152.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111-121.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS ONE*, *9*(2), e89642.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews: Neuroscience*, *9*(11), 856–869. http://doi.org/10.1038/nrn2478

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2012). A meta-Analysis of executive components of working memory. *Cerebral Cortex*, *23*(2), 264–282. http://doi.org/10.1093/cercor/bhs007

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*(6), 527-567.

Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C. A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral Cortex*, *23*(9), 2213-2224.

Okada, B. M., & Slevc, L. R. (in press). 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.* Oxford University Press.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PloS ONE*, *5*(6), e11120.

Palmer, C. (2013). Music performance: Movement and coordination. In D. Deutsch (Ed.), *The Psychology of Music* (3rd ed.) (pp. 405-422). San Diego, CA: Academic Press.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*(4), 374.

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music*, *42*(2), 284–298. http://doi.org/10.1177/0305735612471239

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207.

Rosselli, M., Ardila, A., Santisi, M. N., Arecco, M. D. R., Salvatierra, J., & Conde, A. (2002). Stroop effect in Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, *8*(06), 819-827.

Schellenberg, E. G. (2004). Music Lessons Enhance IQ. *Psychological Science*, *15*(8), 511–514. http://doi.org/10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G. (2006). Long-Term Positive Associations Between Music Lessons and IQ. *Journal of Educational Psychology*, *98*(2), 457–468. http://doi.org/10.1037/0022-0663.98.2.457

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology*, *102*(3), 283-302.

Schellenberg, E. G. (2015). Music training and speech perception: a gene-environment interaction. *Annals of the New York Academy of Sciences, 1337,* 170-177.

Schellenberg, E. G., & Moreno, S. (2009). Music lessons, pitch processing, and g. *Psychology of Music*, 209–221.

Schellenberg, G., & Winner, E. (Eds.) (2011). Music Training and Non-Musical Abilities. *Music Perception*, *29*(2).

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116-

141.

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*(3), 300.

Slevc, L.R., Davey, N., Buschkuehl, M., & Jaeggi, S.M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition, 152*, 199-211.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency does musical ability matter?. *Psychological Science*, *17*(8), 675-681.

Sloboda, J. (1974). The Eye-Hand Span--An Approach to the Study of Sight Reading. *Psychology of Music, 2*(2), 4–10.

Sörqvist, P., Marsh, J. E., & Nöstl, A. (2013). High working memory capacity does not always attenuate distraction: Bayesian evidence in support of the null hypothesis. *Psychonomic Bulletin & Review*, *20*(5), 897-904.

Southgate, D. E., & Roscigno, V. J. (2009). The impact of music on childhood and adolescent achievement*. *Social Science Quarterly*, *90*(1), 4-21.

Strait, D., & Kraus, N. (2011). Playing Music for a Smarter Ear: Cognitive, Perceptual and Neurobiological Evidence. *Music Perception*, *29*(2), 133–146.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

van den Wildenberg, W., van Boxtel, G., van der Molen, M., Bosch, D., Speelman, J., & Brunia, C. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *Journal of Cognitive Neuroscience, 18*(4), 626-636.

Vaughn, K. (2000). Music and Mathematics: Modest Support for the Oft-Claimed Relationship. *Journal of Aesthetic Education*, *34*(3/4), 149–166. http://doi.org/10.2307/3333641

von Bastian, C. C., & Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language, 69*, 36-58.

Werner, P. D., Swope, A. J., & Heide, F. J. (2006). The music experience questionnaire: Development and correlates. *The Journal of Psychology*, *140*(4), 329-345.

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *5*(1), 7-17.

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS ONE*, *9*(6), 1–14. http://doi.org/10.1371/journal.pone.0099868