Measuring the impact of musical learning on cognitive, behavioural and socio-emotional wellbeing development in children

Authors: Rose, D.,[1] Jones Bartoli, A.,[2] & Heaton, P.[2]

[1] Department of Psychology and Sport Sciences, University of Hertfordshire, United Kingdom

[2] Department of Psychology, Goldsmiths College, University of London, United Kingdom

Corresponding Author: d.rose3@herts.ac.uk

Dawn Rose, Alice Jones Bartoli & Pamela Heaton

Abstract

This study investigated the effects of musical instrument learning on the concomitant development of cognitive, behavioural and socio-emotional skills in 38 seven- to nine-year-old children. Pre/post measures of intelligence, memory, socio-emotional behaviour, motor ability and visual-motor integration were compared in children who received either extra-curricular musical training (EMT: *n* = 19) or statutory school music group lesson (SSM: *n* = 19). Results showed a significant association between musical aptitude and intelligence overall. The EMT group showed a significant increase in IQ (7 points), in comparison to 4.3 points for the SSM group, suggesting an effect of musical learning on intelligence. No effects were found for memory, or for visual motor integration or socio-emotional behaviour. However, significant improvements in gross motor ability where revealed for the EMT group only, for the Aiming and Catching composite. With regard to the measure of fluid intelligence, these findings support previous studies (e.g. Forgeard et al., 2008; Hyde et al., 2009; Schellenberg, 2004). The novel use of the Movement Assessment Battery for Children (Henderson, Sugden, & Barnett, 2007) provides evidence that musical learning may support development in a child's ability to judge distance, consider velocity, focus and use their proprioceptive, interoceptive and exteroceptive nervous systems.

Keywords

learning, music education, transfer effects, motor abilities, intelligence, memory, socio-emotional wellbeing

Altenmüller and Schneider described performing music as "the most demanding of human accomplishments" (2008, p. 332). Learning to play music(al instruments) requires multi-modal perception, production and integration skills. These include planning and executing complex motor sequences whilst integrating auditory, visual, tactile and interoceptive, exteroceptive and proprioceptive information in a constant dynamic monitoring mode. Multiple brain regions in both hemispheres and neural networks have been associated with structural and functional changes due to, or concomitant with, musical training (see e.g. Bangert & Altenmüller, 2003; Bengtsson et al., 2005; Imfeld et al., 2009; James et al., 2014; Lee, Chen, & Schlaug, 2002; Mahncke et al., 2006; Oztürk et al., 2008; Schmithorst & Wilke, 2002; Sluming et al., 2002; Stewart et al., 2003). Furthermore, co-activation in subcortical structures such as the basal ganglia and limbic systems suggests musical development may be associated with pleasurable rewards (Herholz & Zatorre, 2012). ~~Consequently~~, Therefore, as music is a whole brain activity, and changes in neural architecture have been observed in line with skill specific learning, the study of expert adult musicians has contributed to the phenomenon of 'metaplasticity' (Wan & Schlaug et al., 2010; Stewart, 2008). It is within this study of metaplasticity that the notion of transfer effects of learning becomes important in terms of childhood development.

However, as research into the development of these skills in children has been undertaken in various ways, comparative understanding of the findings has been complicated. For example, whilst many studies have included longitudinal designs, these have been undertaken with and without the randomisation to different types of interventions (e.g. Moreno et al., 2011; Schellenberg, 2004), or into music and controls only (e.g. Hyde et al., 2009; Putkinen et al., 2014), or participants have been pseudo-randomised[1] (Habibi et al., 2014), or studies have used mixed methods but focused on lived experiences with an emphasis on qualitative data and case studies (McPherson, Davidson, & Faulkner, 2012). This study aimed to explore the concomitant development of cognitive, behavioural and socio-emotional skills during the initial year of musical instrument learning in an ecologically valid setting. Furthermore, the battery of measures was devised to enable consideration of the association between the notion of transfer effects and musical ability.

The term *near transfer* has been used to suggest that musical training increases discrete local skills, such as fine motor ability (see e.g. Costa-Giomi, 1999; 2005; Hyde et al., 2009; Lahav, Saltzman, & Schlaug, 2007; Schlaug et al., 2005). The term *far transfer* has been associated with global abilities, or domain general constructs such as executive function or *g*, measured in terms of IQ (see e.g. Moreno et al., 2011; Schellenberg, 2004; 2006). However, the notion of *transfer effects* and musical learning cannot be considered completely separate from that of aptitude for musical learning (see Ericsson, Krampe, & Tesch-Römer, 1993; Howe, Davidson & Sloboda, 1998). Furthermore, the propensity to learn and/or practice must be supported by opportunity and temperament in order to be realised. Studies have suggested a heritable genetic disposition towards

---

[1] There are convincing ethical reasons for targeting specific groups (see e.g. Costa-Giomi 1999; 2004 and Mehr et al., 2013).

musicality (including the propensity to practice) of up to 70% (Macnamara, Hambrick, & Oswald, 2014; Mosing et al., 2014 a & b; Ullén et al., 2014). Plomin and Deary (2015) described a process of self-selection as an active model of selected environment, in which genotypes can develop into phenotypes. Only one study of five-year-old children has suggested that musical training may increase musical audiation (Flohr, 1981). However, it remains unclear whether there is genetic predisposition towards musical ability, or the ability to practice music, and/or how much phenotypical behaviour is related to enculturation (see e.g. Ericsson, 2007; Gregerson et al., 2013; Hambrick et al., 2014; Johnson, 2011; Turner & Ioannides, 2009; Ukkola-Vuoti et al., 2009; 2013).

Musical enrichment has been associated with an advantage of socio-economic status (SES). That is, increased exposure to music, as well as the provision of musical instrument learning, combines to support the process of enculturation (Hargreaves, 1986; Hallam & Prince, 2003). Alluding to Bourdieu’s concept of cultural capital, Hallam (2010) suggests that a social advantage of musical learning is also likely in terms of the enriched environment provided by parents. This, and the effect of school and relationship with music teachers is an important aspect of developing musical lives, an idea termed *musical transactional regulation* by McPherson et al. (2012) in their exemplary longitudinal study. To encompass a range of experiences, including the potential effects of socio-economic-status (SES), participants for the study were recruitment purposefully to incorporate both state and independent schools where extra-curricular musical instrument lessons are offered, but are either heavily subsidised (state schools) or paid for entirely by parents (independent schools). Data was gathered concerning parents’ levels of education, attitude towards musical learning and on the children’s extra-curricular activities, such as sports clubs, hobbies and arts and crafts as well as music.

Several studies have suggested an association between musical aptitude, learning and cognitive ability (Forgeard et al., 2008; Hyde et al., 2009; Schellenberg, 2004). Furthermore, memory, in particular working memory (WM), is an important cognitive function related to *g* and executive attention (see e.g. Conway, Kane, & Engle, 2003; Kyllonen & Christal, 1990). Meinz and Hambrick (2010) suggested musical training makes a positive contribution of 7.4% in terms of WM capacity, but also suggest that this is highly heritable and domain specific. In terms of children’s studies, some small advantages of musical learning in relation to auditory WM and short-term memory have been reported (Ho et al., 2003; Lee et al., 2007; Rickard et al., 2010; Roden, Kreutz & Bingard, 2012). However, the findings were not equivocal and may be related to the style of musical tuition.

Evidence pertaining to the understanding of how musical instrument learning affects motor and visual skills, and the integration of these (described as visual-motor integration) skills in children has been described as inadequate (McPherson, 2005). Gilbert’s (1980) research (a cross-sectional study of 808 three to six-year-olds) suggested no particular advantage in motor skill development related to musical training. However, Costa-Giomi (1999; 2005) reported a significant increase in performance in measures of motor proficiency in nine-year-old children (*N* = 117) following two years of piano training in a randomised control trial (RCT) in comparison to a control group who were not learning piano, although this advantage (of musical learning) was not observed in the third year. Following Costa-Giomi, developments in neuro-imaging techniques enabled the inclusion of hypotheses related to the notion of pre-existing differences (aligned to musical aptitude).

Forgeard et al., (2008) then reported results of a longitudinal study in which the musically trained children outperformed the control group (of non-musically trained children) in motor learning tasks[2] following 15 months of weekly piano lessons. This was considered evidence of near transfer as these studies as the children were learning piano. The results were not only predicted by duration of training but also associated with evidence of early adaptation observed in the pre-central gyrus, corpus callosum and Heschl's gyrus. Hyde and colleagues (2009) found further evidence that musical training was associated with structural change and better performance on behavioural tasks in six-year-old children. The musically trained group (receiving half hour weekly piano lessons) significantly outperformed the musically untrained control group on a four-finger motor sequencing task (dominant hand). Analyses showed these results predicted neural adaptations, including an increase in grey matter in the right primary auditory cortex, motor areas (such as the precentral gyrus), and in the midbody of the corpus callosum. Conceptually and methodologically it is important to be able separate priming effects, which occurs during encoding and time-on-task, from the concept of near transfer in order to be able to identify which specific skills might contribute either therapeutically and/or educationally. Klingberg (2010) suggested that relying upon a direct measurement of musical skill amounts to a positive bias and does not test for near transfer of ability.

Schlaug and colleagues (2005) have suggested that musical notation training is spatial and the process of learning therefore enhances spatial reasoning. In adults, music notation reading has been correlated with increased grey matter volume and an activation response associated with the temporal cortex in visual-spatial processing tests (Gaser & Schlaug, 2003). Activation in the parietal cortex was also present even when musical notation was simply observed, rather than being performed by musicians (Stewart et al., 2003). Jäncke (2006) suggested that musical notation reading could lead to an increased ability to understand the association between particular visual-spatial shapes and particular sounds and/or musical actions. This is not to say that other musical skills, for example the ability to learn pieces 'by ear' and 'off by heart', are not also exigencies of musical learning that require some visual aspects (for example musicians must be aware of the precise placement of fingers, for example, on a fret board). However, only two studies of children have investigated the development of visual-spatial perception in association with musical learning. Hurwitz et al. (1975) reported a significant increase in visual-motor integration associated with

---

[2] The motor task was a finger tapping paradigm requiring the children to use their index finger to tap the spacebar of a computer keyboard as many times as possible in 20 seconds. This task is performed twice with each hand, beginning with their dominant hand and the scores are averaged (Peters & Durding, 1978, 1979).

musical training, but as their study was cross-sectional they could not claim any level of causality. Furthermore, Hurwitz and colleagues' studied children learning via the Kodály method, which does not require musical notation reading and requires intense levels of parental involvement, known to be an important contributing factor to musical achievement in children (Sloboda & Howe, 1991). Orsmond and Miller (1999) tested 58 children (three and a half to seven years old) before and after four months of Suzuki training, using the Beery Visual Motor Integration test (Beery, 1989). They reported significant improvements and a significant interaction between group, sex and duration of learning for this measure. However, the children in that study had learned via Suzuki method, like the Kodály method, does not use musical notation in the early stages and also includes a high reliance on parental inclusion.

Overall, Mehr and colleagues (2013) have suggested that there is a bias towards reporting the positive effect of musical training. They claimed that in the five RCTs published there is insufficient evidence of transfer effects associating musical training with improved cognitive abilities (such as improved literacy). Therefore, it is appropriate to briefly discuss the co-occurring issues of participant selection and/or randomisation, and motivation (to conduct and participate in) musical studies.

Whilst RCTs offer protection against bias, they cannot necessarily account for motivation to learn, and if this major variable is negated methodologically, how can findings be used in any meaningful way? Costa-Giomi (1999) noted this when discussing her research, an RCT study specifically aimed to provide pianos and lessons to underprivileged children in North America. After finding some positive effects of training after two but not three years, she considered that after an initial surge, enthusiasm for learning and practice waned – only 78 participants (of 117) completed the study. This is related in the literature to ideas regarding the *autotelic value* of practice (Elliott, 1993) and Allport's (1961) concept of *functional autonomy* whereby motive becomes drive, and also builds on the work of Dweck (1986) with regard to motivational behaviours. Dweck persuasively argues that measuring performance on a task in itself does not take into account psychological factors that may influence the outcome. She suggests the move towards a social-cognitive approach of learning has shifted the emphasis towards cognitive mediators such as motivational patterns in terms of goal-orientated behaviours. McPherson, Davidson and Faulkner (2012) further suggest that understanding the nature of musical learning at a fundamental level is imperative, including aspects which can be described as intellectual, creative, social, perceptual and physical. As they convey, musical learning does not take place in isolation. Therefore, whilst this study was quasi-experimental, ecological validity was considered carefully. Children participate in all sorts of activities (such as computer games, cookery, crafts, swimming etc.), which may all contribute to these aspects of development. Therefore, data relating to the amount of hours per week the children took part in activities (including music) was gathered (parent report).

Hypotheses

The first hypothesis for this study is that, in line with Gordon’s assertion, the Primary Measures of Music Audiation (PMMA; Gordon, 1986) will not be correlated with the Wechsler Abbreviated Scale of Intelligence (WASI; Weschler, 1999). The second hypothesis is based on Gordon’s theory that musical aptitude does not stabilise until the age of nine years and is a result of a combination of innate ability and an enriched musical environment (Gordon, 1986). Here the opportunity arises to compare whether extra-curricular musical training has more of an effect on musical audiation than statutory school music. Therefore, H_2 predicts that the extra-curricular music training (EMT) group will increase performance significantly more than the statutory school music (SSM) group on the PMMA over time due to musical training.

The third hypothesis predicts that the EMT group will outperform the SSM group on the overall measure of intelligence as measured in IQ points by the WASI and on measures of auditory WM as measured using the Children’s Memory Scale (CMS). Furthermore, with regard to the specific subtests of the WASI, the EMT group should outperform the SSM group on both vocabulary and matrix reasoning but not on similarities or block design.

With regard to behavioural measures, hypothesis four asserts that he EMT group will perform significantly better than the SSM group on measures of fine and gross motor ability and also visual-motor integration and motor coordination as measured using the Beery VMI and MC (but no differences between groups over time are predicted for the Beery visual perception [VP]).

Finally, hypothesis five predicts that parents and teachers will report higher levels of socio-emotional wellbeing as measured using the Behavioural Assessment System for Children (BASC) in the children in the EMT group in comparison to the SSM group.

Methods and measures

In order to establish whether pre-existing differences related to musicality were apparent, trainable and/or whether musical aptitude was associated with any changes observed in the cognitive measures, the Primary Measure of Musical Audiation (PMMA; Gordon, 1986) was included as a measure specific to this age group. The PMMA is an auditory test in two parts: 40 items of melodic and rhythmic same/different tasks (10 minutes each). They were administered in that order at separate times during the battery. To measure cognitive abilities, standardised tests of intelligence and memory were used. These were Weschler’s Abbreviated Scale of Intelligence (WASI: Weschler, 1999) and the Children’s Memory Scale (CMS; Cohen, 2007). The WASI is a well-known test using various tasks to measure Vocabulary, Similarity (concepts of likeness) for verbal IQ and Block Design and Matrix Reasoning to measure performance IQ. For this study, all four parts were administered to obtain the full scale IQ. The CMS included word lists and digit span tasks, as well as sequencing (such as saying the months of the year backwards). It took approximately 15-20 minutes to administer both the WASI and CMS. To address issues surrounding criticisms of tapping paradigms in children’s musical learning studies (see Sloboda, 2000) a novel measure of fine and gross motor abilities, the Movement Assessment Battery for Children (MABC-2; Henderson, Sugden, & Barnett, 2007) was chosen. This includes measures of Manual Dexterity (such as a pin board task), Aiming and Catching (e.g. throwing a ball against the wall and catching it), and Balance (hopping for example). This took up to 30 minutes to complete. To test the effects of musical learning on visual-motor integration domain, the Beery Visual Motor Integration test (Beery, 2004) was used. The Beery consists of three parts: Visual Motor Integration (VMI; a copying drawing task 10-15 minutes), Visual Perception (a timed shape matching task, 3 minutes), and Motor Coordination (timed and guided drawing tasks, 5 minutes). The Behavioural Assessment System for Children (BASC-2: Reynolds & Kamphaus, 2004) was used to provide an alternative perspective (parents and teachers) from the self-report data associated with socio-emotional wellbeing, in part due to the age of the children. The questionnaires (different for parent/teacher) consist of 150-170 questions about the child’s behaviours, reflecting for example the child’s adaptability. The questions are divided into Clinical and Adaptive groupings. For full descriptions of this battery, see Rose (2016).

The PMMA and Beery were administered in small groups, whilst all other tests were administered individually.

The Research Ethics Committee at Goldsmiths, University of London granted ethical approval for these studies.

Music group was the between-subjects factor and time was the within-subjects factor in the study. Independent *t*-tests or one-way analysis of variance tested group differences at Time 1 (baseline). Repeated measures Analysis of Variance (ANOVA) was used to observe differences over time (main effects) and interactions between groups. Statistical tests were conducted using the Statistical Package for Social Sciences (SPSS; Version 22, IBM Corp.). Planned post hoc analysis (paired samples *t*-tests, or where the assumptions for parametric analyses were not met, Wilcoxon Signed Ranked Tests) were performed to determine change over time for each group. Bonferroni’s method was used to correct *p*-value for multiple comparisons on a measure-by-variable basis. Variance in reported sample sizes is due to missing data. A cut off level of the mean plus or minus three standard deviations was chosen for exclusion based on statistical rather than clinical norms. The size of the participant sample ($N = 38$) is comparable with that of many published studies within the field (see e.g. Fujioka et al., 2006; Norton et al., 2005; Overy, 2003; Schlaug et al., 2009). Due to recruitment limitations, a post-hoc power analysis was performed. For paired samples *t*-tests, this suggested that in order to detect an effect size $d = .3$ (comparable to other studies), with an alpha level of .01, with this sample size there would be a critical value of $t = 2.43$ (power = .29). For these parameters, the effect size would have to reach .55 to achieve a power value above .80. However, when the alpha level was .05 there would be a critical value of $t = 1.69$ amounting to power of .57. To achieve .8 power, the effect size would need to reach .45. To detect a significant effect using bivariate correlation, the Pearson coefficient would need to be $r = .71$ to reach a power value of .80. For repeated measures ANOVA, a partial eta squared value of $\eta_p^2 = .15$ would be equivalent to an effect size of .42 at a power value of .80 for this sample size.

Participants and procedure

Participants were recruited from four schools (two state and two private/independent) from diverse areas across the U.K. One state school specialised in performing arts where the music programme is subsidised, though parents/caregivers were expected to pay £1 towards each lesson.

From the initial 44 participants recruited, 38 completed all measures. The age of the children (21 females) ranged from 85-103 months (from seven years old) and spanned form Year 3 to 4 (U.K.). At Time 1 (baseline) the groups were not statistically significantly different in age ($p = .16$), or handedness ($p = .41$), or sex distribution, χ^2 (1, $N = 38$) = 0.11, $p = .74$.

Half had either self- or parentally-selected to begin extracurricular music lessons amounting to > 1 hour per week over and above statutory music provision, henceforth known as the Extracurricular Music Tuition (EMT) group. The other half, the Statutory School Music (SSM) group received only statutory music group lessons of < 1 hour per week.

The instruments the 19 EMT children reported learning were: seven keyboard/piano, three guitar, two trumpet/horn, one drum kit, and six multiple instruments. Of these six, two were simultaneously learning piano and drums, two were learning both piano and violin, one was learning piano, violin and singing and one was learning piano and guitar.

Music teachers were asked to provide notes on the children's lessons, though only two did. However, the author followed up the students and found 14 students had passed their Grade 1 examinations, or continued to work towards them – and only one student had given up playing entirely (flute). Two students had given up one of their instrument (drum kit and piano), but another had started learning an additional instrument (clarinet). Therefore, whilst individual experiences of teaching and learning cannot be conveyed, the uptake of examinations at least suggests a focus on a performance-based criterion including some understanding of written musical notation.

Results

[INSERT TABLE 1]

Table 1 presents data relates to SES, schools and parental levels of education and investment in musical education for their children. No statistical differences were found between school groups and SES groups when using postcode data[3] (see e.g. Hyndman et al., 1995; Morley et al., 2015; Noble et

---

[3] Postcodes from Jersey, or incomplete postcodes were replaced by a code indicating the parents' level of education and employment as an estimate SES. Classification was in line with the Office of National Statistics system.

al., 2007). However, analysis of the parents' levels of education (Table 2) revealed the EMT group had achieved significantly higher levels, $t(32) = -3.41$, $p = .002$ (equal variances not assumed).

The EMT group parents placed a significantly higher importance on musical learning than those in the SSM group, $t(32) = 2.86$, $p = .008$.

[INSERT TABLE 2]

Table 2 shows the data reported by parents relating to non-academic activities. A statistically significant difference between groups was revealed for musical activity $t(31) = -3.70$, $p = .001$, but not for Leisure Activity $t(29) = -1.43$, $p = .16$ or Physical Activity $t(31) = .18$, $p = .89$.

[INSERT TABLE 3]

Table 3 presents the results from the entire battery of measures.

For the measure of musical aptitude, there was a main effect of time but no interaction between groups. PMMA Composite, $F(1, 35) = 29.94$, $p < .01$, $\eta_p^2 = .46$, and for Tonal, $F(1, 35) = 10.40$, $p < .01$, $\eta_p^2 = .23$, and for Rhythm, $F(1, 35) = 16.54$, $p < .01$, $\eta_p^2 = .32$.

For the measure of g (WASI), there was also a main effect of time $F(1, 35) = 5.67$, $p = .02$, $\eta_p^2 = .14$. Only one subtest, the subtest Matrix Reasoning showed a significant main effect of time, $F(1, 35) = 6.67$, $p = .01$, $\eta_p^2 = .16$. Planned post-hoc analyses revealed a significant increase in IQ points for EMT only, $t(18) = -2.83$, $p = .01$, a mean change of 6.95 points in comparison to the SSM which improved by .17 points on average. The effect size of this change for the EMT group was $d = -.7$ (SSM $d = -.02$).

The only significant effect for the Children's Memory Scale was for the subtest of sequences, in which both groups improved.

There was a main effect of time $F(1, 36) = 5.05$, $p = .03$, $\eta_p^2 = .12$ but no interaction between groups for the total score of the Movement ABC-2. There was also a main effect of time for the subtest Aiming and Catching, $F(1, 35) = 8.94$, $p < .01$, $\eta^2 = .20$. Planned post-hoc analyses revealed a significantly increased performance for the EMT group for Aiming and Catching, $t(18) = -3.51$, $p < .01$, $d = .8$. The SSM group also marginally increased performance, $t(17) = -2.08$, $p = .053$, $d = .2$.

No significant changes were revealed for the Beery tests.

For the BASC, there were no significant changes over time for either group for the teacher and parent composite scores (Behavioural Symptoms Index, Externalising Problems, Internalising Problems, School Problems and Adaptive) or for any of the clinical and adaptive scales. However, some systematic differences were revealed between groups. Teachers reported lower levels of Internalising Problems for the EMT group, $t(25) = 2.61$, $p = .01$, and also for the scale of Anxiety, $t(23.54) = 3.31$, $p < .01$, equal variance not assumed. Parents reported a significant difference between groups at baseline (SSM: $M = 56.33$, $SD = 5.89$; EMT: $M = 49.00$, $SD = 6.18$; $p < .01$), and this continued as a systematic difference over time, $t(27) = 2.49$, $p = .02$, with the EMT group showing significantly lower scores on the Aggression scale.

A significant correlation between the PMMA composite and WASI Full Scale IQ was revealed for the difference between pre and post test $r = .28$, $p = .05$. To correct for measurement error (and therefore avoid Type II errors) this result was disattenuated using Osbourne's (2003) equation revealing a true association, $Tr = .34$, for this relationship. None of the musical and cognitive aptitude subtests were correlated. PMMA Rhythm and MABC-2 Balance were significantly correlated, $Tr = .60$, $p = .02$. The Beery Visual Perception measure was also correlated with the PMMA composite, $Tr = .58$, $p = <.01$.

Discussion

In this study a battery of tests measured cognitive, behavioural and socio-emotional development of the children during their first year of learning their musical instruments. The aim was to investigate the concomitant development of skills that have been associated with musical learning, rather than look at singular concepts of transfer effects. Participants were not randomised to music groups because motivation to choose to learn a musical instrument was seen as important and ethical factors in this (unfunded) study. Socio-economic status, gender and age were equally balanced between groups.

The key findings of this study were the EMT group showed a significant increase in IQ, in particular fluid intelligence, as measured using matrix reasoning. A significant association between musical aptitude and intelligence overall was revealed. Furthermore, significant improvements in gross motor abilities were for the EMT group only (Aiming and Catching composite). However, no effects of the first year of musical learning were found for fine motor skills, memory, or for visual motor integration or socio-emotional behaviour.

Relating the results to the hypotheses, the first investigated Gordon's claim that performance on the PMMA is not related to intelligence (Radocy & Boyle [1979, p. 272] agreed with this claim). However, the results revealed that musical audiation is significantly correlated with intelligence. This supports previous findings suggesting that the PMMA is positively associated (at a low magnitude) with the concept of *g*. Shuter (1968) reviewed 65 studies and found a positive correlation between musical aptitude and intelligence tests ($r = .35$). The disattenuated correlation coefficient found herein concurs ($Tr = .34$), further suggesting that this sample size was adequate to replicate previous findings.

The second hypotheses considered the trainability of audiation. According to Gordon, audiation develops as a result of an enriched musical environment. Though Gordon did not specifically state whether it is trainable or not, he did suggest audiation stabilises around the age of nine years. Here the opportunity arose to compare whether the absolute scores of the EMT group increased more than those of the SSM group. However, whilst the EMT group scored higher on all PMMA scores at baseline, their scores were not significantly higher than the SSM scores. Further, as both groups increased their composite scores significantly (and to a level congruent with 80% power based on the post-hoc power analysis), the results suggest that extra-curricular music training does not necessarily significantly influence PMMA performance over time at this age group. Whilst one explanation of this could be that ceiling effects prevented the EMT scores from increasing at Time 2 relative to Time 1, observation of the histograms for both groups at both pre and post suggest this was not the case. Unfortunately it is not possible to compare this finding with an older sample as the PMMA changes to the IMMA at the age of nine, and is not directly comparable. However, as shown in Flohr (1981) and Schlaug et al., (2005), there is some evidence suggesting the effect of musical training is stronger in younger children than older children perhaps due to differences in neural developmental trajectories (Giedd & Rapoport, 2010).

The third hypotheses predicted that the EMT group would outperform the SSM group on the overall measure of intelligence as measured in IQ points by the WASI as was reported by Schellenberg (2004) findings. Schellenberg reported a seven IQ point increase for his musical group and a four IQ point increase for his control group. The results of this study showed the same IQ point increase for the EMT group but less than one point increase for the SSM control group. However, a 7-point change is less than half of one standard deviation. Whilst the effect size for this change overall (IQ) was very small ($\eta_p^2 = .14$), the effect size for the specific test (matrix reasoning) for the EMT group was large ($d = .6$). Furthermore, the critical t value ($t = 2.43$) for an alpha value of .01 from the post hoc power analysis (< .8) indicates this finding is robust. As no differences were found at Time 1 for either of these factors for age, sex or schools and for SES as inferred by a combination of postcode and parental education levels (analysed as a coded variable) this suggests the in-building of heterogeneity to the research design has been an effective solution against threats to internal validity consummate with quasi-experimental conditions (Mitchell & Jolley, 2012). One further point regarding the significance of the findings presented here is that this advantage is apparent after only one year of training (amounting to approximately 14 hours on average over the duration of one academic year per participant).

Also in relation to hypothesis three, the CMS was used to determine whether EMT in the first year had an effect over and above SSM lessons on measures of auditory memory. According to the CMS (Cohen, 1997), learning and memory are the ability to acquire (new) information, retain and access (stored) information and incorporate the ability to direct and sustain attention/concentration (CMS Manual, Cohen, 1997, p. 11). The Word List subtests of the CMS are divided into Word List Learning (four trials), and Word List Recall, which required the participant to recall the original list (of unrelated words) after a distracter word list has been presented. Our results suggested a trend toward improvement in performance for both groups over time for Word List Learning, ostensibly a simple span measure of auditory STM. For Word List Recall, which requires some consolidation to long-term memory (following a distractor list of different words), there were no statistically significant findings. For the digit span subtests, there were also no significant differences between groups. Further exploration of the effect of musical instrument learning and types of memory (possibly using more thorough *n*-back tasks) will be necessary for future research with regard to transfer effects.

The Sequences task of the CMS requires participants to repeat back a set of semantically grouped either numbers or words (such as the days of the week forwards and backwards under timed conditions) under timed conditions. This task assesses the ability to mentally manipulate and sequence auditory/verbal information as quickly as possible, thereby “placing a heavy demand” on WM (Cohen, 1997, p. 151). Both groups made significant improvements on this subtest. When reporting this finding to the participants’ teachers, none were surprised as the focus during that year (based on the Key Stage 2 Curriculum in the U.K.) was on learning sequences, with specific practice, for example, on learning the months of the year. This may be of interest in light of Klingberg’s assertions related to the concept of transfer. Similarly, behavioural tests were included because in order explore the assumption that the acquisition of a domain specific skill is directly linked to more distant or global skills, measures of near transfer must be distinct from the skills being specifically trained (Klingberg, 2010). These tests were the MABC-2 (Henderson, Sugden, & Barnett, 2007) and the Beery VMI, VP and MC (Beery, 2004). The MABC-2 is a standardised test used to evaluate motor skill in children and adolescents. The measure assesses sensorimotor functioning and motor coordination; specifically focusing on gross motor ability (e.g. jumping, catching), fine motor ability (e.g. drawing, writing) and motor coordination. The test yields a global score (MABC-2 Total) and three component scores, the Manual Dexterity (MD), Balance, and Aiming and Catching (A&C). Changes in gross motor skills associated with musical learning have not been tested before. The use of the MABC-2 is novel within the field. The Beery is designed to assess the extent to which individuals aged between two and eighteen years can integrate their visual and motor abilities (hand-eye coordination). It includes three tests (VMI, VP and MC), but there is no overall score for the Beery.

Regarding the fourth hypothesis, which predicted that the EMT group would perform significantly better than the SSM group on measures of fine and gross motor skills over time as measured using the MABC-2, the analysis of the data for the MABC-2 Total overall measure of motor ability, revealed a main effect of time, with both groups improving, but no significant difference between SSM and EMT groups. For the second level composite measures of the MABC-2, the A&C composite, a significant main effect of time was observed and planned post-hoc analyses revealed a significant improvement in the EMT group. The improvement in the SSM group approached significance level. The composite scores were made up of various tasks, which are subsequently described as third level tasks. The A&C composite included two third level tasks, throwing a beanbag onto a target and throwing and catching a ball. The MD composite included pegboard, sewing and drawing a trail tasks. The Balance component included balancing on a board, hopping and walking along a marked line.

The analysis of the data from the task level for A&C showed that the EMT group outperformed the SSM group on the ‘bean bag’ task of the A&C at the second time point. This task required participants to throw a bean bag onto a marked target across a distance of approximately two metres, requiring hand-eye coordination and judgment regarding velocity, distance and target focus. Whilst the SSM group decreased their performance scores on this task over time, the magnitude of this decrease was not significant. Regarding the ball throwing and catching tasks, there was a main effect of time but no interaction. Planned post-hoc analysis revealed a trend towards significance level in both groups, although this was stronger in the SSM group. As the data for this test were not normally distributed, this was also analysed using non-parametric tests. These revealed that the EMT group improved beanbag throwing over time, whilst SSM group showed no increase in performance on that task. In contrast, the SSM group improved on ball throwing and catching over time whilst no change was observed for the EMT group.

The analysis of the data from the Manual Dexterity component failed to reveal any differences between groups or over time. For the Balance component, as participants achieved ceiling level performance (for the tasks of hopping on each leg and walking along a straight line), these tasks were not analysed.

In summary, whilst the overall analysis of the MABC-2 data showed a significant improvement over time for both groups, the second level composite of Aiming and Catching suggests there were differences between groups. The analysis of the data from the tasks revealed a difference between groups for ball throwing and catching and beanbag throwing onto a marked target. The EMT group outperformed the SSM group on the beanbag task and the SSM group outperformed the EMT group on the ball task. Similarities between these tasks include an understanding of velocity and focused attention, whilst differences between the tasks centre on reaction in order to receive/catch the ball. In Forgeard et al., (2008) and Hyde et al., (2009), a direct effect of piano training on finger tapping and sequences was observed over a slightly longer experimental period. Whilst direct comparison with the finger-tapping paradigm would be invalid due to the mixed instruments learned by the participants in the EMT group, it is noted that no fine motor skills effects were observed in our music-training group. One study which did include gross motor and movement skills (Derri et al., 2001) reported that ten, twice-weekly music and movement interventions resulted in a significantly greater improvement in locomotor skills such as running, jumping and skipping in the experimental group compared with the control group of 68 four to six year olds. Derri and colleagues used the Test of Gross Motor Development (Ulrich, 1985). However, it could be argued that these movements are too far removed from the musical task, or similar to the intervention training, to indicate successful near transfer of acquired skills, which is the focus of this study.

Regarding visual-motor integration (VMI), visual perception (VP) and motor coordination (MC) as measured using the Beery, analyses revealed no significant main effects over time, or interactions between music groups. Failure to replicate the findings of Orsmond & Miller, 1999 was not surprising as their study reported statistical significance based on changes over time in raw, rather than standardised scores, which are now available for the Beery. Hypothesis four specifically predicted that there would be no difference between groups during the experimental period for visual perception. These data support this prediction as measured using the Beery VP.

The final hypothesis considered the development of socio-emotional wellbeing from the perspective of the parents and form teachers (rather than music tutors), rather than using self-report using the BASC-2. Regarding changes in behaviours over the one academic year, neither teachers nor parents reported any significant positive or negative changes for the composites or scales of the BASC-2. However, the data did suggest some systematic differences between groups. Teachers reported that children in the EMT group scored significantly lower than the children in the SSM group for the composite of Internalising Problems, as well as for the clinical scales of Aggression, Anxiety, Conduct Problems, Depression and Hyperactivity. However, once the statistics had been corrected for multiple comparisons, only the scaled of Anxiety remained significantly lower for the EMT group in comparison to the SSM group. The teachers did not report significant systematic differences between EMT and SSM music groups overall for the clinical or adaptive composites or scales. The only scale where parents reported a systematic difference between groups was for Aggression on which the EMT group scored significantly lower than the SSM group. Whilst this finding did not withstand correction for multiple comparisons, the trend towards significance supports other studies suggesting an effect of musical activity in promoting pro-social behaviours (Croom, 2015; Hille & Schupp, 2014; Kirschner & Tomasello, 2010; Moore, Burland, & Davidson, 2003; Rabinowitch, Cross, & Burnard, 2012), though the methodologies and age groups involved in those studies are very different. It is important to note that this sample includes children in mainstream education in the U.K. It was therefore unsurprising that the BASC-2 scores suggested normal behaviours over the experimental period for most of the participants.

In addition to addressing the hypotheses, there are several points raised by this study that require comment. Firstly, the data suggested that the only reported difference between groups was the amount of musical activities they undertook each week and their parent’s positive attitude towards musical learning. These are critical data, not only contributing towards the ecological validity of the study, but also providing further evidence towards the important role of the parents. The context within which musical learning takes place cannot be underestimated as an effective factor as, for example, Davidson et al. (1996), and Sloboda and Davidson (1996) have shown. Further qualitative data can be found in Rose (2016) which supports the importance of the alignment of positive social, biological, and psychological syzygies in the development of musical learning, as described in McPherson et al. (2012).

Limitations

Some of the tests used were, with hindsight, less suitable than others. For example, two tasks on MABC-2 were very easy for the majority of this typically developing sample. Children hit ceiling on the hopping and walking the line tasks, which contribute to the Balance component. Therefore, this composite may not be an accurate measure of change over time. With regard to the Beery tests, scores for the whole sample reduced over time suggesting participants were not attending. Exploration on this matter suggested the children did not find them engaging. For future studies, alternative measures for visual-motor integration and balance should be identified. Finally, it may be that another measure of musicality would have been produced different results. Gordon's PMMA could be described as a measure of short-term musical memory. As was seen in the CMS results, this aspect of cognition was no different between groups. This may be because learning by rote (for example for performance) is not an educational tool utilised in the tradition of western music learning, at least not at this stage (McPherson et al., 2012). In light of the results suggesting musical learning provides an advantage for matrix reasoning (i.e. pattern matching) type tests, perhaps future studies should consider Karma's 2007 Music Test of auditory structuring abilities as this test reports to measure the ability to hear set patterns in sounds.

Conclusion

For this sample, the results suggest that in this sample, individual musical lessons during the first year of learning provide an advantage not only to cognition in terms of fluid intelligence (i.e. problem solving), but also with regard to proprioception (muscular, tendon and joint), exteroception (afferent information pertaining to the mouth, skin and eyes), and possibly towards interoception (concerning the internal organs, such as the inner ear for balance). We argue that the variety of instruments included in the study suggest an overall benefit of understanding the force of pressure exerted causing an effect and being able to temper that accordingly. In musical terms this may be realised as beginning to understand that if one blows, strums or hits to hard or soft, the instrument does not resonate correctly and make the desired sound. Further research may show a direct relationship between the types of skills developed according to instrument, and this had implications in terms of the provision of interventions.

However, the temptation of studying the effects of musical learning with regard to the notion of transfer effects traps us in a false positive in a similar way that beliefs about innate talent did twenty years ago. As McPherson et al. (2012) suggest, "Highly valued outcomes of schooling such as self-discipline, being able to work with others and in a team, problem-solving, and creative thinking are all effectively learned and enhanced through musical participation" (p. 3). It may be that it is these transferable extra-musical skills, more so than transfer effects per se, that provide enduring qualities we can value as part of musical learning, alongside the music itself of course.

References

Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart and Winston.

Altenmüller, E. & Schneider, S. (2008). Planning and performance. In M. Hallam, S. Cross, & I. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 332–343). Oxford, UK: Oxford University Press.

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience*, *4*(1), 26, 1-14. http://doi.org/10.1186/1471-2202-4-26

Beery, K. E. (1989). *The developmental test of visual–motor integration* (3rd ed.). Cleveland, OH: Modern Curriculum Press.

Beery, K. E. (2004). *The Beery-Buktenica developmental test of visual-motor integration*. Minneapolis, MN: NCS Pearson.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, *8*(9), 1148–1150. http://doi.org/10.1038/nn1516

Cohen, M. (1997). *Children's memory scale (CMS)*. San Antonio, TX: Psychological Corporation.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. http://doi.org/10.1016/j.tics.2003.10.005

Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, *47*(3), 198. http://doi.org/10.2307/3345779

Costa-Giomi, E., Flowers, P. J., & Sasaki, W. (2005). Piano lessons of beginning students who persist or drop out: Teacher behavior, student behavior, and lesson progress. *Journal of Research in Music Education*, *53*(3), 234–247. http://doi.org/10.1177/002242940505300305

Croom, A. M. (2014). Music practice and participation for psychological well-being: A review of how music influences positive emotion, engagement, relationships, meaning, and accomplishment. *Musicae Scientiae*, *19*(1), 44–64. http://doi.org/10.1177/1029864914561709

Davidson, J. W., Howe, M. J., Moore, D. G., & Sloboda, J. A. (1996). The role of parental influences in the development of musical performance. *British Journal of Developmental Psychology*, *14*(4), 399-412.

Derri, V., Tsapakidou, A., Zachopoulou, E., & Kioumourtzoglou, E. (2001). Effect of a music and movement programme on development of locomotor skills by children 4 to 6 years of age. *European Journal of Physical Education*, *6*(1), 16–25. http://doi.org/10.1080/1740898010060103

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. http://doi.org/10.1037/0003-066X.41.10.1040

Elliott, D. J. (1993). On the values of music and music education. *Philosophy of Music Education Review*, *1*(2), 81–93.

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology*, *38*(1), 4–34.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Flohr, J. W. (1981). Short-term music instruction and young children's developmental music aptitude. *Journal of Research in Music Education*, *29*(3), 219. http://doi.org/10.2307/3344995

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PloS One*, *3*(10), e3566. http://doi.org/10.1371/journal.pone.0003566

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., &, & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*(10), 2593–2608. http://doi.org/10.1093/brain/awl247

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240–9245. Retrieved from http://www.jneurosci.org/content/23/27/9240.short

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron*, *67*(5), 728–34. http://doi.org/10.1016/j.neuron.2010.08.040

Gilbert, J. (1980). An assessment of motor music skill development in young children. *Journal of Research in Music Education*, *28*(3), 167. http://doi.org/10.2307/3345234

Gordon, E. E. (1986). *Manual for the Primary Measures of Music Audiation and the Intermediate Measures of Music Audiation.* Chicago, IL: G.I.A.

Gregersen, P. K., Kowalsky, E., Lee, A., Baron-Cohen, S., Fisher, S. E., Asher, J. E., … Li, W. (2013). Absolute pitch exhibits phenotypic and genetic overlap with synaesthesia. *Human Molecular Genetics*, *22*(10), 2097–2104. http://doi.org/10.1093/hmg/ddt059

Habibi, A., Ilari, B., Crimi, K., Metke, M., Kaplan, J. T., Joshi, A. A., … Damasio, H. (2014). An equal start: absence of group differences in cognitive, social, and neural measures prior to music or sports training in children. *Frontiers in Human Neuroscience*, *8*, 690. http://doi.org/10.3389/fnhum.2014.00690

Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education*, *28*(3), 269–289. http://doi.org/10.1177/0255761410370658

Hallam, S., & Prince, V. (2003). Conceptions of musical ability. *Research Studies in Music Education*, *20*(1), 2–22. http://doi.org/10.1177/1321103X030200010101

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45. http://doi.org/10.1016/j.intell.2013.04.001

Hargreaves, D. J. (1986). *The developmental psychology of music*. Cambridge, UK: Cambridge University Press.

Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement Assessment Battery for Children, 2nd Edition (Movement ABC-2)*. London: The Psychological Corporation.

Herholz, S., & Zatorre, R. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron*, *76*(3), 486–502.

Hille, A., Arnold, A., & Schupp, J. (2014). Leisure behavior of young people: Education-oriented activities becoming increasingly prevalent. *DIW Economic Bulletin*, *4*(1), 26–31. Retrieved from http://www.econstor.eu/handle/10419/91599

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.

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*(3), 399–407. http://doi.org/10.1017/S0140525X9800123X

Hurwitz, I., Wolff, P. H., Bortnick, B. D., & Kokas, K. (1975). Nonmusical effects of the Kodaly music curriculum in primary grade children. *Journal of Learning Disabilities*, *8*(3), 167–174. http://doi.org/10.1177/002221947500800310

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *Journal of Neuroscience*, *29*(10), 3019–3025. http://doi.org/10.1523/JNEUROSCI.5118-08.2009

Hyndman, J. C. G., Holman, C., D'Arcy, J., Hockey, R. L., Donovan, R. J., Corti, B., & Rivera, J. (1995). Misclassification of social disadvantage based on geographical areas: Comparison of postcode and collector's district analyses. *International Journal of Epidemiology*, *24*(1), 165–176. http://doi.org/10.1093/ije/24.1.165

IBM. (2013). Statistical Package for Social Sciences (SPSS). Armonk, NY: IBM Corp.

Imfeld, A., Oechslin, M. S., Meyer, M., Loenneker, T., & Jancke, L. (2009). White matter plasticity in the corticospinal tract of musicians: A diffusion tensor imaging study. *NeuroImage*, *46*(3), 600–607. http://doi.org/10.1016/j.neuroimage.2009.02.025

James, C. E., Oechslin, M. S., Van De Ville, D., Hauert, C.-A., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure & Function*, *219*(1), 353–366. http://doi.org/10.1007/s00429-013-0504-z

Jäncke, L. (2006). From cognition to action. In J. Alternmuller, E., Weisendamnger, M., Kesselring (Eds.). *Music, motor control and the brain* (2nd ed., pp25-38). Oxford, UK: Oxford University Press.

Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. http://doi.org/10.1016/j.dcn.2010.07.003

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

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317–324. http://doi.org/10.1016/j.tics.2010.05.002

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. http://doi.org/10.1016/S0160-2896(05)80012-1

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience*, *27*(2), 308–14. http://doi.org/10.1523/JNEUROSCI.4822-06.2007

Lee, D. J., Chen, Y., & Schlaug, G. (2002). Corpus callosum: Musician and gender effects. *Neuroreport*, *14*(2), 205–209. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=14657319

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, *17*(3), 336–344. http://doi.org/10.1016/j.learninstruc.2007.02.010

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*(8), 1608–1618. http://doi.org/10.1177/0956797614535810

Mahncke, H. W., Bronstone, A., & Merzenich, M. M. (2006). Brain plasticity and functional losses in the aged: scientific bases for a novel intervention. *Progress in Brain Research*, *157*, 81–109. http://doi.org/10.1016/S0079-6123(06)57006-2

McPherson, G. E. (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music*, *33*(1), 5–35. http://doi.org/10.1177/0305735605048012

McPherson, G. E., Davidson, J. W., & Faulkner, R. (2012). *Music in our lives: Rethinking musical ability, development and identity*. Oxford University Press.

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), e82007. 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

Mitchell, M., & Jolley, J. (2012). *Research design explained*. Belmont, CA: Cengage Learning.

Moore, D. G., Burland, K., & Davidson, J. W. (2003). The social context of musical success: A developmental account. *British Journal of Psychology*, *94*(4), 529–549. http://doi.org/10.1348/000712603322503088

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), 1425–1433. http://doi.org/10.1177/0956797611416999

Morley, D., Till, K., Ogilvie, P., & Turner, G. (2015). Influences of gender and socioeconomic status on the motor proficiency of children in the UK. *Human Movement Science*, *44*, 150–156. http://doi.org/10.1016/j.humov.2015.08.022

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*(9), 1795–1803. http://doi.org/10.1177/0956797614541990

Mosing, M. A., Pedersen, N. L., Madison, G., & Ullén, F. (2014). Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. *PloS One*, *9*(11), e113874. http://doi.org/10.1371/journal.pone.0113874

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Exley, S., Barnes, H., & Dibben, C. (2007). *The English indices of deprivation*. London: Department of Communities and Local Government.

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., & Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain and Cognition*, *59*(2), 124–134.

Orsmond, G. I., & Miller, L. K. (1999). Cognitive, musical and environmental correlates of early music instruction. *Psychology of Music*, *27*(1), 18–37. http://doi.org/10.1177/0305735699271003

Overy, K. (2003). Dyslexia and music. *Annals of the New York Academy of Sciences*, *999*(1), 497–505. http://doi.org/10.1196/annals.1284.060

Öztürk, A. H., Tasçioglu, B., Aktekin, M., Kurtoglu, Z., & Erden, I. (2008). Morphometric comparison of the human corpus callosum in professional musicians and non-musicians by using in vivo magnetic resonance imaging. *Journal of Neuroradiology*, *29*(1), 29–34. Retrieved from http://www.em-consulte.com/en/article/119679

Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, *20*(1), 98–108. http://doi.org/10.1038/mp.2014.105

Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N., & Huotilainen, M. (2014). Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiology of Learning and Memory*, *110*, 8–15. http://doi.org/10.1016/j.nlm.2014.01.007

Rabinowitch, T.-C., Cross, I., & Burnard, P. (2012). Long-term musical group interaction has a positive influence on empathy in children. *Psychology of Music*, *41*(4), 484–498. http://doi.org/10.1177/0305735612440609

Radocy, R. E., & Boyle, J. D. (1979). *Psychological foundations of musical behavior.* Springfield, IL: Charles C. Thomas.

Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavioral Assessment System for Children manual.* (2nd ed.). Circle Pines, MN: AGS.

Rickard, N. S., Vasquez, J. T., Murphy, F., Gill, A., & Toukhsati, S. R. (2010). Benefits of a classroom based instrumental music program on verbal memory of primary school children: A longitudinal study. *Australian Journal of Music Education*, *1*(1), 36–47.

Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. *Frontiers in Psychology*, *3*, 572. http://doi.org/10.3389/fpsyg.2012.00572

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.

Schlaug, G., Forgeard, M., Zhu, L., Norton, A., Norton, A., & Winner, E. (2009). Training-induced neuroplasticity in young children. *Annals of the New York Academy of Sciences*, *1169*, 205–208. http://doi.org/10.1111/j.1749-6632.2009.04842.x

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, *1060*, 219–230. http://doi.org/10.1196/annals.1360.015

Schmithorst, V. J., & Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: A diffusion tensor imaging study. *Neuroscience Letters*, *321*(1–2), 57–60. http://doi.org/10.1016/S0304-3940(02)00054-X

Shuter, R. (1968). *The psychology of musical ability.* London: Methuen.

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, *4*(10), 397–403. http://doi.org/10.1016/S1364-6613(00)01531-X

Sloboda, J. A., & Howe, M. J. A. (1991). Biographical precursors of musical excellence: An interview study. *Psychology of Music*, *19*(1), 3–21. http://doi.org/10.1177/0305735691191001

Sloboda, J. A., & Davidson, J. (1996). The young performing musician. In I. Deliége & J. A. Sloboda (Eds.) *Musical beginnings: Origins and development of musical competence* (pp. 171-190). New York: Oxford University Press.

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's Area in male symphony orchestra musicians. *NeuroImage*, *17*(3), 1613–1622. http://doi.org/10.1006/nimg.2002.1288

Stewart, L. (2008). Fractionating the musical mind: insights from congenital amusia. *Current Opinion in Neurobiology*, *18*(2), 127–130. http://doi.org/10.1016/j.conb.2008.07.008

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *NeuroImage*, *20*(1), 71–83. http://doi.org/10.1016/S1053-8119(03)00248-9

Turner, R., & Ioannides, A. A.. (2009). Brain, music, and musicality: Inferences from neuroimaging. In S. M. & C. Trevarthen. (Ed.), *Communicative musicality: Exploring the basis of human companionship* (pp. 147–184). Oxford, UK: Oxford University Press.

Ukkola-Vuoti, L., Kanduri, C., Oikkonen, J., Buck, G., Blancher, C., Raijas, P., … Järvelä, I. (2013). Genome-wide copy number variation analysis in extended families and unrelated individuals characterized for musical aptitude and creativity in music. *PloS One*, *8*(2), e56356. http://doi.org/10.1371/journal.pone.0056356

Ukkola, L. T., Onkamo, P., Raijas, P., Karma, K., & Järvelä, I. (2009). Musical aptitude is associated with AVPR1A-haplotypes. *PloS One*, *4*(5), e5534. http://doi.org/10.1371/journal.pone.0005534

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. http://doi.org/10.1016/j.paid.2014.01.057

Ulrich, D. A. (1985). *Test of gross motor development.* Austin, TX: Pro-Ed.

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist*, *16*(5), 566–577. http://doi.org/10.1177/1073858410377805

Weschler, D. (1999). *Weschler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.

Table 1. Parental descriptive data

		Group	
		SSM	EMT
Descriptor	Level	Frequency	Frequency
SES Level	High	8	11
	Middle	4	5
	Low	7	3
School Type	State	7	7
	Private/Independent	12	12
Level of Parental Education[a]	*O level/GCSE	8	5
	**A level/Highers	6	1
	Undergraduate Degree Level	2	4
	Post Graduate Degree Level	0	8
Attitude to Musical Learning[b]	Essential	2	5
	Important	3	10
	A pleasant past time	10	3
	Little or no value	1	0

[a,b] Missing data

* O level/GSCE are the equivalent to the U.S. High School Diploma

** A level (England)/Highers (Scotland) are the university entry qualifications in the U.K., equivalent to a High School Diploma plus Advanced Placement in the U.S., or the International Baccalaureate.

Table 2. Total hours of weekly activity participation as reported by parents

Music Group		Hours per Week Physical Activity	Hours per Week Musical Activity	Hours per Week Leisure Activity
SSM (*n* = 15)	Mean (SD)	5.07 (2.28)	1.53 (1.06)	3.33 (1.54)
	Median	5.00	1.00	4.00
	Minimum	2.00	.00	.00
	Maximum	11.00	4.00	6.00
EMT (*n* = 18)	Mean (SD)	4.94 (1.63)	3.22 (1.48)	4.06 (1.29)
	Median	5.00	3.00	4.00
	Minimum	2.00	1.00	.00
	Maximum	8.00	6.00	6.00

Table 3. Results of Battery of Measures

Measure	SSM		Statistic	EMT		Statistic
	Pre Mean (SD)	Post Mean (SD)	Group Change Over Time	Pre Mean (SD)	Post Mean (SD)	Group Change Over Time *Significant without need for correction ** Significant after correcting for multiple comparisons.
Primary Measure of Musical Aptitude - (Percentiles)						
Composite	52.47 (18.48)	73.32 (12.01)	$t(18) = -4.76$, $p < .01$, $d = -1$**	61.71 (22.81)	77.17 (18.46)	$t(17) = -3.03$, $p < .01$, $d = -.7$*
Tonal	51.68 (19.90)	67.84 (15.88)	$t(18) = -3.17$, $p = .01$, $d = -.7$**	64.37 (22.98)	73.89 (20.62)	ns
Rhythm	52.68 (21.82)	73.05 (13.31)	$t(18) = -3.82$, $p < .01$, $d = -.9$**	57.42 (25.63)	72.72 (17.73)	$t(17) = -2.13$, $p =.048$, $d = -.5$**
Weschler's Abbreviated Scale of Intelligence						
FSIQ	103.83 (11.71)	104.00 (9.02)	ns	108.84 (15.39)	115.79 (14.69)	$t(18) = -2.83$, $p =.01$, $d = -.7$*
VIQ^b	107.22 (15.54)	109.39 (14.31)	ns	114.89 (20.20)	122.37 (18.17)	ns
PIQ^a	101.11 (19.16)	100.17 (12.50)	ns	103.42 (18.52)	109.16 (15.83)	$t(18) = -2.57$, $p = .02$, $d = -.6$**
Vocabulary^a	50.33 (6.84)	53.94 (10.37)	ns	52.95 (15.46)	60.42 (12.05)	$t(18) = -2.31$, $p = .03$, $d = -.5$
Similarities^a	56.17 (11.36)	56.00 (5.66)	ns	61.58 (10.23)	61.95 (9.82)	ns
Block Design^a	50.94 (9.99)	47.39 (8.20)	ns	51.68 (11.65)	52.74 (8.68)	ns
Matrix Reasoning^a	50.89 (11.08)	52.78 (7.67)	ns	51.79 (10.01)	56.42 (8.29)	$t(18) = -2.81$, $p = .01$, $d = -.6$**
Children's Memory Scale^b						
Word List Learning (aSTM)	10.32 (2.91)	11.05 (3.21)	ns	9.11 (3.40)	10.63 (3.22)	ns
Word List Recall (aLTM)	12.37 (3.04)	11.37 (2.83)	ns	10.53 (3.20)	11.58 (3.15)	ns
Digit Span Forwards (aSTM)	11.26 (2.79)	12.32 (2.69)	ns	10.53 (3.20)	11.58 (3.15)	ns
Digit Span	10.11	10.74	ns	9.32	10.16	ns

Backwards (aWM)	(2.71)	(2.66)		(2.16)	(2.99)	
Sequences	10.00	12.37	*t*(18) = -4.51, *p* < .01, *d* = -.1**	10.53	12.89	*t*(18) = -4.61, *p* < .01, *d* = -.1**
(Attention)	(2.21)	(2.14)		(3.55)	(3.75)	
Movement Assessment Battery for Children – 2[b]						
Total Score	9.53	10.26	ns	9.26	10.47	ns
	(2.67)	(2.45)		(1.94)	(2.30)	
Manual Dexterity	8.53	9.26	ns	9.05	9.32	ns
	(3.27)	(3.45)		(2.66)	(2.85)	
Aiming and	9.67	10.44	ns	10.16	10.84	*t*(18) = -3.51, *p* < .01, *d* = .8**
Catching	(2.83)	(2.75)		(2.59)	(2.09)	
Balance	10.37	11.58	ns	10.84	11.37	ns
	(2.06)	(2.80)		(2.09)	(2.67)	
Beery[b]						
Visual-Motor	101.95	98.74	ns	105.95	104.24	ns
Integration	(12.92)	(7.44)		(13.01)	(14.67)	
Visual Perception	99.11	96.94	ns	103.11	99.86	ns
	(16.34)	(17.96)		(15.47)	(12.02)	
Motor Control	87.42	86.33	ns	90.63	91.11	ns
	(6.41)	(12.52)		(13.53)	(16.31)	

[a] T Scores, [b] Standardised Scores