International Journal of Educational Research 89 (2018) 10–21

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

# A feasibility study of the impact of the Kodály-inspired music programme on the developmental outcomes of four to five year olds in England

Beng Huat See[a,*], Lindsay Ibbotson[b,1]

[a] School of Education, Durham University, United Kingdom
[b] Kodály Music Specialist – Early Years, United Kingdom

ARTICLE INFO

*Keywords:*
Kodály inspired music
Early Years
Randomized controlled trial

ABSTRACT

The Kodaly approach to music is a fun and interactive way to introduce music to young children. There is currently some evidence suggesting that this approach to music training can have beneficial effects on children. No rigorous studies, however, have been conducted to test its effects on very young children. This paper presents the results of a pilot study investigating the impact of the Kodály approach to music on the developmental outcomes of 56 pre-school children in one school in the North East of England. Children were individually randomized to either the intervention or to a delayed control group. Progress was measured using the Early Learning Goals (ELGs) set out in the National Curriculum. Data collected across the ELGs over two terms indicate that on almost all measures the treatment children outperformed control children in terms of social-emotional and behavioural development (ES = +0.71 for behaviour; ES = 0.32 for self-confidence; ES = 0.43 for relations). No effects were observed for writing after one term (ES = 0). Small improvements were observed after two terms (ES = +0.4). For reading, intervention children performed worse than control after one term (ES = −0.16), but after two terms they caught up with the control children (ES = +0.03). Progress in number skills was maintained after two terms (ES = +0.3). The small sample weakens the conclusion that we can make, but the results suggest promise. The trial also shows that teachers can be effectively trained to deliver the programme.

## 1. Introduction

There has been an increasing interest in the potential benefits of music on pupil achievement at school. In the UK there have already been movements to support music in primary and secondary schools. A number of political and sector-led initiatives, (such as Sing Up; Musical Futures; Inspire-Music), supported by successive governments have seen an increased focus on music in schools. Following the Henley Review (Henley, 2011), the National Plan for Music Education was published (DfE, 2011) to ensure a high quality music education, where children from all backgrounds and every part of England would have the opportunity to learn a musical instrument, make music with others through whole-class ensemble teaching programmes for a minimum of a term, learn to sing, and have clear progression routes available. Pupils from disadvantaged backgrounds could potentially be offered music lessons, using the Pupil Premium Fund, though this is at individual school discretion.

* Corresponding author.
*E-mail addresses:* b.h.see@durham.ac.uk (B.H. See), lindsay.ibbotson@hotmail.co.uk (L. Ibbotson).
[1] Clints, Marske, Richmond, DL11 7LY.

https://doi.org/10.1016/j.ijer.2018.03.002
Received 11 February 2018; Received in revised form 28 February 2018; Accepted 3 March 2018
0883-0355/ 

Despite the many initiatives and policies, which were largely focused on the older primary and secondary school pupils there were still persistent barriers to access, according to an Institute of Education report (Creech, Saunders, & Welch, 2016). It was suggested that these barriers could be traced back to a lack of opportunity to engage with progressive musical activity at a far earlier stage in the educational experience of the child.

Research suggests that to be effective music training needs to be introduced at an early age, preferably before the age of seven, where neurological changes are more likely to occur (e.g. Musacchia, Sams., Skoe, & Kraus, 2007; Penhune, 2011; Steele, Bailey, Zatorre, & Penhune, 2013; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Wan & Schlaug, 2010). A review of arts education (See & Kokotsaki, 2015) found that relatively few studies were conducted for children under the age of five. Although the Kodály approach to music is already practised in a number of schools for very young children in the Early Years settings, as far as we know, there have been no rigorous randomized controlled trials conducted to test its impact − at least not in the UK. The report recommended pilot trials of the Orff, Kindermusik and Kodaly approach to music on young children. It is against this background that we conducted this pilot trial with Reception year children (age 4–5) using a randomized controlled (RCT) design.

In England Early Years education refers to the provision of care and education for children from birth up to the age of five. Reception year is the final year in the Early Years Foundation Stage of Education. All schools and registered Early Years providers have to comply with the Early Years Foundation Stage Framework, which sets the standards for the learning, development and care of children. Children in the Early Years are assessed on their development across 17 Early Learning Goals (ELG). These include communication and language, physical, social, emotional and behavioural outcomes, literacy and numeracy as well as creativity. A more detailed explanation is provided in the Methods section.

This paper presents the findings from this small pilot. The programme was piloted in one school in the Tees Valley in the North East of England. The rationale for conducting the research in this area and for targeting the Reception age children is as follows:

Currently a number of school research projects are concentrated in London and the South East. Few of the schools in the high poverty areas in the north have been similarly funded, least of all for music education research. For this reason we targeted this project here.

Children in Reception year are at an age when they are still young enough to benefit most from early exposure to music education. There has been little robust experimental research conducted to test the impact of music on very young children. This project aimed to address that gap.

One of the persistent barriers to the provision of music education at an early age is the lack of training for generalist primary school teachers and Early Years practitioners. This project integrated staff training within the intervention.

## 2. Background evidence

There is indicative evidence that the Kodály, Kindermusik, (which is based on Kodály), and Orff methods of learning music can have beneficial effects on the cognitive development of young children (Duncan, 2007; Hetland, 2000; François, Chobert, Besson, & Schön, 2013; Myant, Armstrong, & Healy, 2008; Schellenberg, 2004). Positive effects were reported for a range of outcomes: creativity (Barkóczi & Pleh, 1982; Duncan, 2007), spatial-temporal ability (Gromko & Poorman 1998; Hetland, 2000; Hurwitz, Wolff, Bortnick, & Kokas, 1975), IQ scores (Kaviani, Mirbaha, Pournaseh, & Sagan, 2014), reading and language (Myant et al., 2008) and social skills (e.g. Ghasemtabar, Hosseini, Arab, Naghashian, & Poudineh, 2015). Some studies also suggest that the Kodály approach to learning improves psychomotor skills, perception ability as well as performance in other academic areas such as reading and maths (DeVries, 2001).

However, most of these studies were not randomised controlled trials. Hetland's (2000) review, for example, claimed to be a *meta*-analysis of experimental studies, but only five of the 15 studies randomised participants to treatment conditions. And almost all the studies in the review involved very small numbers of children (with an average of 47 cases in each study). Six studies had a sample of between five and 70. The rest was much smaller. Barkóczi and Pleh (1982) also claimed to use experimental design controlling for family background, but the comparison groups were randomly selected rather than randomly allocated. For example, three classes were randomly selected − one class came from a Kodály music primary school, the other two classes were taken from a local primary school, one of which was "selected by chance" to receive special music lessons, while the other class which formed the control received the normal music lessons. It is therefore possible that there may be inherent or unobservable differences between the groups of children being compared which may have accounted for the differences in outcomes. Duncan's (2007) study compared children whose parents volunteered to take part in the study with those whose parents did not. Participants were therefore self-selected. There were only 32 children in this study. The few studies that involved randomising participants to treatment groups were all very small scale. Gromko and Poorman, (1998) and Ghasemtabar et al. (2015) had only 30 children in their study. It was also unclear if the children in Ghasemtabar et al.'s study were randomly assigned. Moreover, where randomisation was used, classes or schools rather than individuals were randomised. This reduces the statistical power of detecting effects (if any) since the number of cases compared would be the classes/schools rather than the individuals.

Because of weaknesses in these studies (e.g. lack of comparison groups and non-random allocation of comparison groups), the lack of replication and inconsistent findings across studies, these findings have to be interpreted with caution. More robust and rigorous evaluations are needed to confirm the causal links. This would involve randomization of participants into control and treatment, having a pre- and post-test comparison.

Although the Education Endowment Foundation (EEF) had recently funded two randomized controlled trials evaluating the impact of music on academic achievement (Rhythm for Reading and Act, Sing and Play), the results were not conclusive. The Act, Sing and Play programme (Haywood et al., 2015) was delivered to primary school age children and showed no evidence of positive

gains for participating children (ES for maths was + 0.003 and ES for literacy was +0.03). The study compared children doing music with children doing drama. There was no inactive control group. Since music and drama are both arts activities, the study was unable to provide evidence of the impact of arts or music. It was also unclear whether these teachers understood the effective, progressive musical sequencing in the teaching − what to teach first − which is a key feature of the Kodály approach. As stated in the report, there were inconsistencies in the implementation, and the less experienced teachers needed more guidance. Studies have indicated that the quality of teaching is an important factor in ensuring effective outcomes (Hallam, 2015). Hallam also noted in her book *The Power of Music* that a positive interpersonal relationships between participants and the person delivering the music activity is essential in ensuring success.

The Rhythm for Reading programme (Styles, Clarkson, & Fowler, 2014) involves children taking part in rhythm-based exercises like clapping, stamping and chanting while reading musical notation. It shares many elements in common with the – Kodály approach to music. Results showed a very small effect on children's reading ability (ES = +0.03). Slightly bigger effects were noted for children eligible for free school meals (ES = +0.11), but there was no beneficial effect for the low attainers. It has to be noted that Rhythm for Reading was originally developed for primary school children, but in the EEF-funded trial the programme was delivered to first year secondary school pupils. This may explain the small effects. The study also reported that poor behaviour and lack of engagement may have hindered effective implementation. Process evaluation suggests that the programme might be more suitable for younger children.

In summary while there have been some attempts to test the effects of music in general, no robust large-scale studies have been conducted on the Kodály approach to music on very young children. Research in this area tended to be small-scale, correlational or used convenience samples with non-random allocation of participants (e.g. matched comparison groups). There was therefore an urgent need for more robust research using a randomized controlled design.

## 3. The intervention

The Kodály approach to music, developed by Hungarian composer Zoltán Kodály, *(after visiting England and learning of the work of Sarah Glover and John Curwen)*, is based on the teaching, learning and understanding of music through the experience of singing and musical games. Children can learn basic music skills through their voice and body − there is no particular need to invest in other instruments or equipment.

The Kodály approach to music education is child centred, playful and physical, and taught in a logical, sequential manner. It involves children's active participation in the activities. This sequential learning process follows the natural developmental pattern used in learning a language, which is, aural, written, and then read.

- Aural − oral − kinesthetic
- Written − pictoral − abstract
- Read − recognized

The logical progression of learning moves in small steps from the simple to the complex. Songs begin with a limited vocal range, with opportunities for singing, movement for learning, and plenty of repetition during the games. There are 3 stages in the learning process:

- **Preparation** − the unconscious learning stage, which lasts for many weeks, depending on pupils age and experience. Children learn the repertoire by listening, watching, then copying the singing and movements.
- **Presentation** − as Kodály said, "Children learn best what they already know" − when the children are ready, the teacher chooses some of the repertoire to make an aspect of the learning conscious, introduce musical vocabulary, and where appropriate, introduce a symbol, (e.g. the symbol for a one beat sound).
- **Practice** − using and reinforcing the learning in many ways. Children will sing familiar songs and find the new concept within them, take on challenges, make judgements, listen and appraise, improvise and compose.

Each session in the trial includes work on 4 concepts simultaneously: pulse, rhythm, pitch and structure, though mostly at the preparation stage until the children are ready.

This approach is believed to be very effective with young children who learn the musical elements spontaneously through playing, singing of musical games and songs. This is believed to help develop awareness of pitch, pulse, rhythm, space and time, and the experience of active contribution to the group, thus laying the foundations for the skills of musical memory, simple notation, creative improvisation, and, not least, social interaction. The intervention involves the use of a variety of some popular children's songs and rhymes, such as *Hot Cross Buns, This Old Man,* and *Mrs White had a Fright* and optional simple musical instruments. Children are given the opportunity to explore how sounds can be changed, sing simple songs from memory, and play musical games.

For this pilot, the programme initially ran for one term, and involved two Reception year classes in one school. The sessions were delivered by the same music specialist with Kodály training on a daily basis, running from Tuesday to Friday each week, over 10 weeks. Each session ran from 9.30am for about 15 min. This was after the children had arrived, and settled in their classroom for 'carpet time'. They would then split into their two groups, and the music group would move into a large room on the upper floor of the school building for their session, while the other children stayed in one of the two Reception classrooms for their normal curriculum activities. One of the regular classroom teachers accompanied the intervention group and stayed with them during the

| Activities and resources | Mediating factor | Change mechanism | Outcomes: Impact |
|---|---|---|---|
| • Daily 15 min session first delivered by music specialist, and observed by class teacher<br>• Then delivery of lesson by class teacher<br>• Simple props like hand puppets, sticks, claves and rag dolls<br>• Structured lesson plans with a list of songs developed by the Kodály music specialist | • Teacher engagement with the intervention<br>• Active participation of children in music activities (counting, tapping, clapping and singing)<br>• Regular practice of actions and songs<br>• Positive interpersonal relationship between deliverer and children<br>• Consistent delivery of sessions<br>• Sessions are highly interactive, enjoyable, fun and inclusive | • Increased enjoyment of lessons<br>• Increased motivation<br>• Better listening skills<br>• Understand turn-taking<br>• Increased attention/focus<br>• Develop coordination (e.g. hand clapping, keeping in tempo)<br>• Understand counting in sequence<br>• Better memory (remembering lyrics to the songs) | • Improved psycho-motor skills<br>• Improved self-confidence<br>• More focused and better attention<br>• Improved social interaction and cooperation<br>• Improved number skills<br>• Improved literacy skills<br>• Improved phonological awareness<br>• Positive attitude towards learning |

| Assumptions | External Factors |
|---|---|
| Support from school leadership<br>Learners have 15 minutes of Kodaly sessions for two terms<br>Teachers received training from trained specialists<br>Teachers practise what they have been trained to do<br>Teachers are engaging and fun | Learner absences<br>Staff changes and staff absences |

**Fig. 1.** The logic model.

session. The teacher joined in the activities and observed the trained-specialist deliver the programme thus receiving on-the-job training. To avoid diffusion teachers were told to avoid using these activities in the first term in their regular classes.

To test the effect of a longer duration, the intervention group continued for another term. In the second term the control children joined the intervention children in their normal class format. This means they were no longer in randomized groups. The intervention ran for eight weeks in this term. The class teacher who had attended the specialist sessions in the first term now took up the delivery with her re-integrated class, and the specialist continued working with the other re-integrated class. The songs and games already known by the intervention group were shared with the children who had previously been in the control group.

The logic model in Fig. 1 shows how the activities can lead to improvements in children's cognitive and non-cognitive outcomes and the change mechanism that is essential to developing these skills.

## 4. Relevance

This research is particularly relevant and timely now with the recent call for music and arts education to be given priority in UK schools. The introduction of the English Baccalaureate has seen a decline in interest and takeup in arts at GCSE. Concerns are now raised that these education reforms could stifle creativity and innovation. The House of Lords, for example, argued for arts to be part of the core curriculum to encourage the development of creativity, critical thinking, motivation and self-confidence – skills necessary for innovation. Such skills are also believed to help children learn academic skills. Michael Leigh, the Oscar-award winning director, said it was ridiculous to think of arts as the preserve of the privileged, and that:

"Art should be a core subject of all subjects, like English is, but even more so."

However, it is music in particular that is being highlighted by neurological research as an activity that promotes development of logical thinking (e.g. Sanders, 2012), creativity and expression (Passanisi, Di Nuovo, Urgese, & Pirrone, 2015). Funding bodies like the Education Endowment Foundation in the UK are now funding trials to test the benefits of arts education on young people's learning and other softer outcomes (https://educationendowmentfoundation.org.uk/public/files/Application_Guidance/CulturalLearning-ApplicationGuidanceNotes.pdf)

As the Kodály approach does not require any special equipment beyond the participants themselves, and only a basic understanding of the concepts of beat, rhythm and pitch, it presents a simple and effective way of introducing music to young children,

without budgetary constraints. And since the outcomes of the music programme are in line with the Early Learning Goals set out in the EYFS (Early Years Foundation Stage) guidelines, this approach potentially can help children meet the learning and development goals expected of them. The findings of this study will therefore be of interest to:

- Head Teachers and school leaders to consider making space for the adoption of a simple method of music training
- Teacher trainers − developing much needed provision for Early Years and Primary Music training for student teachers
- Music Educators
- Policy makers

## 5. Aims

The main aim of the study was to assess the impact of a Kodály-inspired music education on the developmental outcomes of children in the Reception year (age 4–5). The secondary aims of the pilot were to find out if one term of delivery was adequate, and whether regular classroom teachers could be trained on the job to teach the programme. In the UK the music curriculum is normally delivered by the generalist teacher (i.e. a regular class teacher) rather than a specialist music teacher.

The main research questions are therefore:

- Does the Kodály approach to music training enhance pre-school children's developmental outcomes (literacy, numeracy, social, emotional and behavioural outcomes) as identified in the Early Years Foundation Stage Framework?
- Does having an additional term of intervention improve outcomes?
- Can regular classroom teachers be effectively trained on the job to teach the programme?

However, as it was a pilot study and the first RCT on such a programme we also wanted to use the opportunity to see if it was feasible to randomise children individually across two classes, specifically looking into the time-tabling constraints and the logistics of mixing children from the two classes.

## 6. Sample

The sample was taken from one state primary school in Teesside, located in the North East of England. Fifty-six children from two Reception year classes (age 4–5) were individually randomized to either receive the intervention immediately (n = 28) or business-as-usual control (n = 28) using a waiting-list design. Because of the relatively small number of pupils involved, individual randomisation was preferred as it increases the power to detect effects (if any). Randomising by class would mean that effectively we would have only two cases instead of 56. Individual randomisation also has the advantage of reducing teacher effects since both groups would include children from the two classes. Any differences in the children's outcomes therefore could not be solely attributed to differences in teacher effectiveness. In this pilot the two classes had a common time-table which made it possible for individual randomisation across the two classes.

To avoid demoralization and also for ethical reasons (so that control children were not excluded from the programme), control children were given the opportunity to take part in the music activities after one term. This means that in the second term of delivery there was no clean control group. But this allowed us to compare children who were exposed to two terms of activity with those who had only one term of exposure.

The school provided the names of the children for randomisation, which was carried out by an independent evaluator from Durham University using graphpad (http://www.graphpad.com/quickcalcs/randomize2/).

### 6.1. Characteristics of the school

The trial school is a voluntary aided school with a pupil population of 376. It is a fairly large primary school compared to the national average (Table 1). Table 1 shows the characteristic of the trial school compared to the national average for primary schools in England. The school was rated Outstanding in the 2011 Ofsted inspection (the last time it was inspected). It is a high performing school with 83% of pupil achieving the expected level at Key Stage 2 (end of primary school assessment) compared to only 61% for the national average (Table 1). The school also has below average proportion of disadvantaged children.

**Table 1**
School characteristic.

| Variable | National average | Trial school |
|---|---|---|
| **Size of schools** | 201–300 | 376 |
| **Proportion of boys** | 51.3% | 50.8% |
| **Proportion meeting expected standard in reading, writing and maths** | 61% | 83% |
| **Proportion of pupils eligible for Free School Meals (FSM) at anytime in the last 6 years** | 24.9% | 2% |
| **Proportion of pupils with Special Educational Needs (SEN) support** | 2.9% | 0.3% |
| **Proportion of pupils for whom Eng is an additional (EAL)** | 20.8% | 4.6% |

Although the developer had plans to work in a wider context with schools located in areas of economic deprivation, this feasibility study was trialled in a school that was particularly supportive and keen to participate, even in the absence of funding. This only partially reflects the aim to carry out the research in an area of economic deprivation, but was nevertheless, an acceptable starting point.

## 7. Methods

The impact evaluation was a simple two-armed randomized controlled design where children from two Reception year classes were individually randomized to either receive the intervention immediately or to business-as-usual control. In addition we thought it would be interesting to also see how pupils respond to the programme. So we asked the class teacher to provide three case study observations on individual children, selected from three groups − one whose learning outcomes were judged to be not yet reaching the level expected at the end of Reception Year (Pupil A); one whose learning outcomes were as 'expected' (Pupil B), and one child whose learning outcomes 'exceeded' the level expected (Pupil C). These were based on the teacher's professional judgement, and so potentially subjective.

### *7.1. Outcomes*

Children's developmental outcomes were assessed using the national standards of assessment known as the Early Years Foundation Stage (EYFS) framework (https://www.gov.uk/government/publications/eyfs-profile-exemplication-materials).

Within this framework children's developmental progress is assessed in 7 areas. The level of progress children are expected to attain at the end of the EYFS is defined by the Early Learning Goals (ELG). For each of the 7 developmental areas there are a number of learning goals. The 7 developmental areas and Early Learning Goals are:

**Communication and language**
ELG 1: listening and attention
ELG 2: understanding
ELG 3: Speaking
**Physical development**
ELG 4: moving and handling
ELG 5: health and self-care
**Personal, social and emotional development**
ELG6: self-confidence and self-awareness
ELG7: managing feelings and behaviour
ELG 8: making relationships
**Literacy**
ELG 9: reading
ELG 10: writing
**Mathematics**
ELG 11: numbers
ELG 12: shape, space and measures
**Understanding the world**
ELG 13: people and communities
ELG 14: the world
ELG 15: technology
**Expressive arts and design**
ELG 16: exploring and using media and materials
ELG 17: being imaginative.

The ELG assessments are based on teachers' observations of the child in their day-to-day interactions, samples of children's work, photographs and contributions from parents. These assessments are externally moderated at the local authority level to ensure accuracy of practitioner judgements. The learning outcomes are judged based on three levels:

- meeting the level of development expected at the end of the reception year (expected)
- exceeding this level (exceeding) or
- not yet reaching this level (emerging)

The use of standardized commercial tests, which are independent of teacher judgements, was explored but the final decision was to use the ELG assessment for a number of reasons. Most commercially produced tests for under five year olds involve one-to-one delivery. This would incur additional costs. They are also very time-consuming. Since the project was conducted without funding, this was not deemed feasible. The alternative was to use the Early Learning Goals assessment as it offered a number of advantages. First it reduced the burden of testing both for the school and the children since it is an assessment which the school was already using anyway. An added advantage was that it would help to see if the intervention was able to address the learning needs of the children as identified in the national curriculum.

#### 7.1.1. Primary outcomes

The primary outcomes of interest were children's cognitive skills: literacy and mathematics performance. These are measured using the teacher assessed Early Learning Goals (ELG) 9 (reading), ELG 10 (writing), ELG 11 (numbers) and ELG 12 (shapes, space and measures)

#### 7.1.2. Secondary outcomes

The secondary outcomes include children's personal, social and emotional development, and creativity. These are measured using children's performance on the following learning areas: self-confidence and self-awareness (ELG 6), behaviour (ELG 7), making relationships (ELG 8) and imagination (ELG 17).

These outcomes were identified by the programme developer as areas on which they thought the intervention would have the most impact. Children's developmental levels are assessed at different time points in the year.

The gradings by the teachers were by subjective teacher assessment of levels achieved in 4 subject areas (literacy; mathematics; personal, social and emotional development; and expressive arts and design). For each level of achievement we have allocated a number score. For example, if a child achieved what might be expected for a 22–36 month old child, but not yet begun to attempt what might be expected for the next band of development for the 30–50 month old child, a number score 1 is awarded. The higher the score, the better is the child's performance for their age.

1 Achieved 22–36 not started 30–50
2 Achieved 22–36, working towards 30–50
3 Started 30–50 and 40–60 but not achieved
4 Achieved 30–50 Not started 40–60
5 Achieved 30–50 Started 40–60
6 Achieved 40–60 working towards ELG
7 ELG working towards Exceeding

In this pilot study, we were concerned only with literacy (reading and writing), mathematics (numbers, shapes, space and measures), personal, social and emotional development (self-confidence, self-awareness, managing behaviour and making relationships) and expressive arts and design (being imaginative).

### 7.2. Analysis

The two groups were not balanced at baseline, with the control group slightly ahead of the treatment group for most of the measures. Therefore, the gain scores were used instead of the post-test scores. Impact was measured by comparing the progress or gain scores between the two groups. This is then converted to Hedges' g effect size using the pooled standard deviation. This is calculated as the difference between grades for treatment and control children. The difference between the mean gain scores for the treatment and control groups divided by the pooled standard deviation is the effect size, which is the size of the difference between the groups.

Comparisons were made first of children's outcomes after one term of delivery (between autumn and spring data) and then after two terms of delivery (between autumn and summer data). The assessment for the autumn term formed the baseline scores. This was when the children first started Reception class. We then collected formative ELG assessments after the Easter holidays, that is, at the end of the Spring Term. This formed the comparative data for the end of the first (Spring) term of intervention. The final summative ELG data were collected after the assessment on the 20th of June, following the 8 week second term (Summer) intervention with the re-integrated classes.

We do not report significant tests and confidence intervals as these are misleading and not appropriate. Significant tests are tests that tell us the probability of observing the results we get assuming that there is no difference between the groups. But the answer that we really want is whether there is a difference between the groups given the results that we have. Unfortunately significant tests do not give us the answer to the latter question. This is a misinterpretation of significant tests (Colquoun, 2014; Perezgonzalez, 2015; Gorard, 2016; Pharoah, Jones, & Siddartha, 2017).

Instead we calculate the Number Needed to Disturb (NNTD), which is defined as the number of counterfactual cases needed to alter the finding. By comparing the number of missing cases to the number of counterfactual cases needed to disturb (NNTD) the finding we can determine whether the number of missing cases is large enough to alter/explain the findings. It is a measure of how stable the result is after attrition (an alternative test of sensitivity). NNTD is calculated as the effect size multiplied by the number of cases in the smallest group in the comparison.

## 8. Ethical concerns

Parents were informed by the school about the programme, but parental consent was not sought as the programme was seen as part of the curriculum which the school was doing anyway.

## 9. Results

All pupil data were available. No children dropped out of the trial. This section reports first on the results after one term of delivery and then results of two terms of delivery.

### 9.1. Impact on academic outcomes after one term of delivery

Tables 2–5 show that in terms of cognitive development, intervention children made bigger gains than control children in number skills (ES = +0.35) and spatial concepts (ES = +0.04) after one term of exposure to the intervention. However, no gains were made for literacy (reading and writing). For education interventions an effect size of 0.2 is considered typical.

**Table 2**
Comparison of mean scores after one term of delivery (reading).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 26.07 | 2.43 | 28.04 | 2.236 | 1.97 | 2.37 | |
| **Control** | 28 | 25.96 | 2.90 | 28.25 | 1.956 | 2.29 | 1.61 | |
| **Total** | 56 | 26.02 | 2.65 | 28.14 | 2.084 | | 2.01 | −0.16 |

**Table 3**
Comparison of mean scores after one term of delivery (writing).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 8.07 | 1.82 | 11.04 | 1.67 | 2.96 | 1.21 | |
| **Control** | 28 | 8.57 | 2.22 | 11.55 | 1.753 | 2.96 | 1.53 | |
| **Total** | 56 | 8.32 | 2.028 | 11.29 | 1.713 | 2.96 | 1.36 | 0 |

**Table 4**
Comparison of mean scores after one term of delivery (numbers).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 35.29 | 3.75 | 40.07 | 0.94 | 4.79 | 3.65 | |
| **Control** | 28 | 36.21 | 2.04 | 40 | 1.05 | 3.79 | 1.77 | |
| **Total** | 56 | 35.75 | 3.03 | 40.04 | 0.99 | 4.29 | 2.88 | +0.35 |

**Table 5**
Comparison of mean scores after one term of delivery (shapes, space and measures).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 24.68 | 1.61 | 28.5 | 1.67 | 3.82 | 1.70 | |
| **Control** | 28 | 25 | 1.98 | 28.75 | 1.71 | 3.75 | 1.53 | |
| **Total** | 56 | 24.84 | 1.79 | 28.63 | 1.68 | 3.79 | 1.60 | +0.04 |

### 9.2. Impact on personal, social, emotional and creative outcomes after one term of delivery

Tables 6–9 show the impact of the intervention on children's social, emotional, behavioural development and on creativity. The results suggest that the Kodály approach to learning music has beneficial effects on children's social, emotional and behavioural development after one term of exposure. Intervention children made the biggest improvements in their self-confidence (ES = +0.42), behaviour (ES = +0.56) and social relationships (ES = +0.47) compared to the control children.

**Table 6**
Comparison of mean scores after one term of delivery (self-confidence and self-awareness).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 18.11 | 0.69 | 18.79 | 0.42 | 0.68 | 0.55 | |
| **Control** | 28 | 18.39 | 0.69 | 18.82 | 0.39 | 0.43 | 0.63 | |
| **Total** | 56 | 18.25 | 0.69 | 18.8 | 0.40 | 0.55 | 0.60 | +0.42 |

**Table 7**
Comparison of mean scores after one term of delivery (managing feelings & behaviour).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 26.54 | 0.922 | 27.75 | 0.52 | 1.21 | 0.83 | |
| **Control** | 28 | 27.18 | 0.77 | 27.93 | 0.26 | 0.75 | 0.75 | |
| **Total** | 56 | 26.86 | 0.90 | 27.84 | 0.42 | 0.98 | 0.82 | +0.56 |

**Table 8**
Comparison of mean scores after one term of delivery (making relationships).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 23.64 | 1.06 | 24.79 | 0.42 | 1.07 | 0.9 | |
| **Control** | 28 | 24.07 | 0.77 | 24.71 | 0.46 | 0.71 | 0.54 | |
| **Total** | 56 | 23.86 | 0.94 | 24.75 | 0.44 | 0.89 | 0.76 | +0.47 |

**Table 9**
Comparison of mean scores after one term of delivery (imagination).

| | N | Mean scores autumn | SD | Mean scores spring | SD | Gain scores | SD | 'Effect' size |
|---|---|---|---|---|---|---|---|---|
| **Treatment** | 28 | 17.79 | 1.81 | 19.89 | 1.197 | 2.11 | 1.449 | |
| **Control** | 28 | 18.5 | 1.262 | 20.29 | 0.6 | 1.79 | 1.197 | |
| **Total** | 56 | 18.14 | 1.59 | 20.09 | 0.95 | 1.95 | 1.33 | +0.24 |

### *9.3. Impact of children's cognitive, social, emotional and behavioural development after two terms of delivery*

Table 10 shows the results after two terms of exposure to the programme. Although the control children are also now exposed to the programme, they had only one term of exposure. The results suggest that the positive effects of the programme are maintained after two terms. What is interesting is that intervention children have now made bigger improvements in literacy and spatial concepts compared to control children. This suggests perhaps that at least a two- term delivery is needed in order for effects on literacy (reading and writing) to be realized.

Since both control and treatment children were now involved in the music activities, it is safe to say that there are no effects of teacher expectation (i.e. a situation where knowledge of treatment allocation may bias teachers' assessment of children's progress).

The results are rather mixed, but it shows that the positive effects on the initial treatment children have been maintained in the second term (Table 11). On almost all measures the treatment children continued to outperform control children. Only for self-confidence and making relationships did the gap close. The effect sizes are slightly smaller suggesting that perhaps the control children have made bigger improvements once exposed to the programme. There is a suggestion that the social interaction of children in their normal classes may have enhanced the impact as bourne out by the data gathered after the re-integration of the normal class groups (Hallam, 2015).

To test how secure the finding is, we calculated the Numbers Needed to Disturb, which is the effect size multiplied by the number of cases in the smallest group. For Number skills, NNTD is 9.8 (0.35 × 28). This means that it will take approximately 10 missing cases or missing data to alter the findings. Since there were no missing data/cases, the finding can therefore be considered very secure. In other words, the results cannot be caused by missing data.

**Table 10**
Comparisons of mean gains after two terms of delivery.

| | Control | | | Treatment | | | Total | | | Effect size |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | N | SD | Mean | N | SD | Mean | N | SD | |
| Self-confidence | 0.61 | 28 | 0.685 | 0.82 | 28 | 0.612 | 0.71 | 56 | 0.65 | +0.32 |
| Behaviour | 0.82 | 28 | 0.772 | 1.46 | 28 | 0.922 | 1.14 | 56 | 0.90 | +0.71 |
| Relations | 0.89 | 28 | 0.737 | 1.29 | 28 | 1.049 | 1.09 | 56 | 0.92 | +0.43 |
| Reading | 3.57 | 28 | 2.588 | 3.64 | 28 | 2.512 | 3.61 | 56 | 2.53 | +0.03 |
| Writing | 3.96 | 28 | 1.953 | 4.71 | 28 | 1.718 | 4.34 | 56 | 1.86 | +0.4 |
| Number | 4.79 | 28 | 2.043 | 5.71 | 28 | 3.75 | 5.25 | 56 | 3.03 | +0.3 |
| Shape | 5 | 28 | 1.981 | 5.32 | 28 | 1.611 | 5.16 | 56 | 1.79 | +0.18 |
| Imagination | 2.5 | 28 | 1.262 | 3.21 | 28 | 1.813 | 2.86 | 56 | 1.59 | +0.45 |

**Table 11**
Comparing effects after one term and two terms of exposure.

| Outcomes | Effect size after 1 term of delivery | Effect size after 2 terms of delivery |
|---|---|---|
| **ELG6 – SELF CONFIDENCE** | +0.42 | +0.32 |
| **ELG7 – BEHAVIOUR** | +0.56 | +0.71 |
| **ELG8 – RELATIONS** | +0.47 | +0.43 |
| **ELG9 – READING** | −0.16 | +0.03 |
| **ELG10 – WRITING** | 0 | +0.4 |
| **ELG11 – NUMBER** | +0.35 | +0.3 |
| **ELG12 – SHAPE** | +0.04 | +0.18 |
| **ELG17 – IMAGINATION** | +0.24 | +0.45 |

For details about progress made in each of the Early Learning Goals, see Appendix A Tables A1 and A2.

## 10. Discussion

The results of this pilot experiment suggest that the Kodály musical activity can have beneficial effects on children's developmental outcomes, which include both cognitive and non-cognitive skills. Improvements in number skills and spatial concepts (cognitive abilities) were apparent after one term of delivery. Effects on literacy (reading and writing) were only realized after two terms of delivery. Children continued to make progress after two terms. This suggests that at least two terms of delivery are needed for impact to be realized, especially for literacy. Whether the effects will be maintained after the intervention stops cannot be ascertained as this is beyond the scope of the project.

The biggest impacts were on social, emotional and behavioural outcomes (non-cognitive skills). The effects for cognitive skills, especially for reading are small suggesting that literacy may be more difficult to shift, but the improvement from an effect size of −0.16 in the first term to an effect of +0.03, although small, is a positive sign. This finding supports the justification for at least a two-term delivery for greater effects, especially for literacy.

Although the effects for cognitive skills may be small, it has to be mentioned that they are typical of randomized controlled trials of education interventions. Big effect sizes are not uncommon in correlational studies or studies using matched comparisons, or simple pre-post comparisons with no comparators. One should not confuse the effects of a proper randomized trial with those of weaker designs.

While it is possible that growth in children's developmental skills is to be expected as a result of maturation, the case is made for the individual randomization. This would ensure that both groups are equal at the outset in terms of their characteristics and backgrounds as well as any unobservable characteristics. As explained earlier, individual randomization also takes account of differences in teacher effectiveness. Therefore, if improvements in confidence and behaviour were the result of maturation or even teacher effectiveness, both groups should make the same progress. The fact that they did not suggests that the improvements could be attributed to the music intervention.

However, because of the very small sample and the reliance on data from only one school, the results have to be interpreted with caution. The school is above the national average in terms of KS2 outcomes and has below the national average proportion of disadvantaged children, so similar results may not be replicated in other schools. But overall, the results are promising and encouraging. This gives justification for a bigger trial to be conducted. Future research could consider a wider range of schools with higher proportion of disadvantaged children to test its effects on different types of schools in different geographical locality.

Another weakness of the study is the use of teacher assessment, which some may argue is not reliable as teachers were not blind to the randomization. To a certain extent biased judgements are minimized as these assessments are moderated externally to ensure accuracy of assessment. Neverthess, future trials could look at using standardized tests for a more objective measure of the outcomes. We are pleased to say that at the time of writing, the Education Endowment Foundation (EEF) in partnership with the Royal Society of Arts has funded a national study as part of the Learning about Culture programme to test the impact of the Kodály approach on children's academic outcomes, targeting 1,800 five to six year olds in 60 schools.

This trial has demonstrated that while it is feasible to randomise individual children within school, research also suggests that the social interactions of children in their normal friendship group (Hallam, 2015) could enhance the impact. Future research may want to consider the impact of class randomization vs individual randomization.

The support from the headteacher was key to ensuring the successful implementation of the trial. The headteacher was instrumental in ensuring that the two classes had a common time-table. This made individual randomization of children in the two classes possible. For successful implementation of such a programme, the support of the school leadership is crucial. This ensures that adequate time is given for the delivery, and space available for the activities.

The trial has also shown that teachers could easily pick up the skills in delivering the programme simply by observing and participating in the activities. It was observed that after one term the class teacher was able take the class independently.

The class teacher reported that children enjoyed the lessons and were actively involved in the activities. Some children who were initially shy and reticent about participating eventually joined in and were soon singing and clapping along. It was observed that the child who was assessed as performing below the expected level (Pupil A) had shown marked improvements in confidence over the two terms. When she first arrived at Reception year she was initially quiet and would not join in with the other children during class Circle Time. Since taking part in the music intervention she had become more confident and even volunteered to sing solo in front of a

class of 30. Her class teacher also commented on how confident she had become, putting up her hands to answer questions in class, and playing with other children. Pupil B (the child assessed to have met the expected level), although confident in his own friendship group, was rather shy with adults, often refusing to join in the music activities. However, since starting the programme, he has grown to be more confident in front of adults. He even volunteered to sing solo, something which his teacher said he would never have done before. He was now enthusiastically looking forward to the session often asking when they were going to have music lessons. Pupil C, a high ability child, on the other hand, was very confident and enthusiastic, frequently volunteering answers and sometimes leading the sessions. But he had the habit of shouting our answers unsolicited. Towards the end of the programme, it was observed that he had begun to understand the concept of turn taking and listening to other children's answers. Although these are all positive outcomes, we cannot say for sure if they could be attributed to the programme as no similar observations were made with the control children. Such progress with the children could be the result of maturation. So although such evidence may be limiting on its own, when combined with the impact evaluation it suggests that the programme has some promise in developing children's social, emotional and behavioural skills.

In summary, this trial has shown that the programme can benefit young children in raising their self-confidence and improving their social and behavioural outcomes. Effects on the cognitive domains (literacy and numeracy), however, are small and are only observed after two terms of delivery, suggesting that at least two terms of delivery are necessary for effects to be realised.

## Appendix A

**Table A1**
Comparison of gain scores after one term of delivery.

| Outcomes | Control | | | Treatment | | | Effect size |
|---|---|---|---|---|---|---|---|
| | Mean | N | Std. Deviation | Mean | N | Std. Deviation | |
| ELG1 Listening | 0.46 | 28 | 0.576 | 0.32 | 28 | 0.476 | −0.27 |
| ELG2 Understanding | 1 | 28 | 0.72 | 0.89 | 28 | 0.685 | −0.16 |
| ELG3 Speaking | 1.46 | 28 | 0.793 | 1.86 | 28 | 1.433 | +0.34 |
| ELG4 Moving & handling | 3.82 | 28 | 1.467 | 3.71 | 28 | 1.584 | −0.1 |
| ELG5 Health & self-care | 2.07 | 28 | 0.663 | 1.79 | 28 | 1.371 | −0.26 |
| ELG6 Self-confidence& self-awareness | 0.43 | 28 | 0.634 | 0.68 | 28 | 0.548 | +0.42 |
| ELG7 Managing feelings & behaviour | 0.75 | 28 | 0.752 | 1.21 | 28 | 0.833 | +0.56 |
| ELG 8Making relationships | 0.71 | 28 | 0.535 | 1.07 | 28 | 0.9 | +0.47 |
| ELG9 Reading | 2.29 | 28 | 1.607 | 1.96 | 28 | 2.365 | −0.16 |
| ELG10 Writing | 2.96 | 28 | 1.527 | 2.96 | 28 | 1.201 | 0 |
| ELG11 Numbers | 3.79 | 28 | 1.771 | 4.79 | 28 | 3.645 | +0.35 |
| ELG12 Shape, space and measures | 3.75 | 28 | 1.531 | 3.82 | 28 | 1.701 | +0.04 |
| ELG13 People & communities | 0.36 | 28 | 0.731 | 0.39 | 28 | 0.737 | +0.04 |
| ELG14 The world | 0.79 | 28 | 1.134 | 0.71 | 28 | 0.897 | −0.08 |
| ELG15 Technology | 0.36 | 28 | 0.488 | 0.64 | 28 | 0.621 | +0.49 |
| ELG16 Exploring media & materials | 4.82 | 28 | 1.657 | 4.04 | 28 | 1.71 | −0.46 |
| ELG17 Being imaginative | 1.79 | 28 | 1.197 | 2.11 | 28 | 1.449 | +0.24 |

**Table A2**
Comparison of mean gain scores for each of the 17 Early Learning Goals after 2 terms of delivery.

| | Control | | | Treatment | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | N | Std. Deviation | Mean | N | Std. Deviation | Mean | N | Std. Deviation |
| ELGGains1 | 0.5 | 28 | 0.577 | 0.43 | 28 | 0.69 | 0.46 | 56 | 0.631 |
| ELGGains2 | 1.46 | 28 | 0.637 | 1.29 | 28 | 0.897 | 1.38 | 56 | 0.776 |
| ELGGains3 | 2.14 | 28 | 1.239 | 2.68 | 28 | 2.091 | 2.41 | 56 | 1.724 |
| ELGGains4 | 4.79 | 28 | 1.572 | 5 | 28 | 2.568 | 4.89 | 56 | 2.112 |
| ELGGains5 | 3.82 | 28 | 0.983 | 3.64 | 28 | 1.283 | 3.73 | 56 | 1.136 |
| ELGGains6 | 0.61 | 28 | 0.685 | 0.82 | 28 | 0.612 | 0.71 | 56 | 0.653 |
| ELGGains7 | 0.82 | 28 | 0.772 | 1.46 | 28 | 0.922 | 1.14 | 56 | 0.903 |
| ELGGains8 | 0.89 | 28 | 0.737 | 1.29 | 28 | 1.049 | 1.09 | 56 | 0.92 |
| ELGGains9 | 3.57 | 28 | 2.588 | 3.64 | 28 | 2.512 | 3.61 | 56 | 2.528 |
| ELGGains10 | 3.96 | 28 | 1.953 | 4.71 | 28 | 1.718 | 4.34 | 56 | 1.861 |
| ELGGains11 | 4.79 | 28 | 2.043 | 5.71 | 28 | 3.75 | 5.25 | 56 | 3.029 |
| ELGGains12 | 5 | 28 | 1.981 | 5.32 | 28 | 1.611 | 5.16 | 56 | 1.797 |
| ELGGains13 | 0.36 | 28 | 0.731 | 0.5 | 28 | 1.106 | 0.43 | 56 | 0.931 |
| ELGGAins14 | 0.89 | 28 | 1.197 | 0.86 | 28 | 1.044 | 0.88 | 56 | 1.113 |
| ELGGains15 | 0.5 | 28 | 0.694 | 0.82 | 28 | 0.772 | 0.66 | 56 | 0.745 |
| ELGGains16 | 6.07 | 28 | 2.308 | 6.68 | 28 | 3.031 | 6.38 | 56 | 2.687 |
| ELGGAins17 | 2.5 | 28 | 1.262 | 3.21 | 28 | 1.813 | 2.86 | 56 | 1.589 |

## References


Barkóczi, I., & Pleh, C. (1982). *Music makes a difference.* Zoltán Kodály Pedagogical Institute of Music: Budapest.

Colquoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*. http://dx.doi.org/10.1098/rsos.140216http://rsos.royalsocietypublishing.org/content/1/3/140216.

Creech, A., Saunders, J., & Welch, G. (2016). *Musical pride: Music education in plural communities. Report for 'The Mix' Luton Leicester Slough.* London: UCL, Institute of Education.

DeVries, P. (2001). Reevaluating Common Kodály practices. *Music Educators Journal, 88*(3), 24–27.

DfE (2011). *The importance of music: The national plan for music education.* London: Department for Education [DFE 00086-2011].

Duncan, D. J. (2007). *The relationship between creativity and the Kindermusik experience. Unpublished Master of Science thesis.* Missouri: University of Central Missouri.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex, 23*(9), 2038–2043.

Ghasemtabar, S. N., Hosseini, M., Arab, S., Naghashian, H., & Poudineh, Z. (2015). Music therapy: An effective approach in improving social skills of children with autism. *Advanced Biomedical Research, 27*(4), 157.

Gorard, S. (2016). Damaging real lives through obstinacy: Re-emphasising why significance testing is wrong. *Sociological Research On-line, 21*, 1. http://www.socresonline.org.uk/21/1/2.html.

Gromko, J. E., & Poorman, A. S. (1998). The effect of music training on preschoolers' spatial-temporal task performance. *Journal of Research in Music Education, 46*(2), 173–181.

Hallam, H. (2015). *The power of music: A research synthesis on the impact of actively making music on the intellectual, social and personal development of children and young people.* London: International Music Education Research Centre.

Haywood, S., Griggs, J., Lloyd, C., Morris, S., Kiss, Z., & Skipp, A. (2015). *Creative futures: Act, Sing and Play. Evaluation report and executive summary*London: Education Endowment Foundation.

Henley, D. (2011). *Music education in England: A review for the Department of Education and the Department for Culture, Music and Sport.* London: Department for Education.

Hetland, L. (2000). *The relationship between music and spatial processes: A meta-analysis. Unpublished EdD thesis.* Cambridge, MA: Harvard University.

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://dx.doi.org/10.1177/002221947500800310.

Kaviani, H., Mirbaha, H., Pournaseh, M., & Sagan, O. (2014). Can music lessons increase the performance of preschool children in IQ tests? *Cognitive Processing, 15*(1), 77–84.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the USA, 104*(40), 15894–15898.

Myant, M., Armstrong, W., & Healy, N. (2008). Can music make a difference? A small scale longitudinal study into the effects of music instruction in nursery on later reading ability. *Educational and Child Psychology, 25*(3), 83–100.

Passanisi, A., Di Nuovo, S., Urgese, L., & Pirrone, C. (2015). The influence of musical expression on creativity and interpersonal relationships in children. *Procedia − Social and Behavioural Sciences, 191*, 2476–2480.

Penhune, V. B. (2011). Sensitive periods in human development: Evidence from musical training. *Cortex, 47*, 1126–1137.

Perezgonzalez, J. D. (2015). Null hypothesis significant tests. A mix-up of two different theories: The basis for widespread confusion and numerous misinterpretations. *Scientometrics, 102*(1), 411–432.

Pharoah, P. D. P., Jones, M. R., & Siddartha, K. (2017). *P-values and confidence intervals: Not fit for purpose. bioRxiv preprint posted online 24 August 2017.* https://www.biorxiv.org/content/biorxiv/early/2017/08/24/180117.full.pdf.

Sanders, E. (2012). Investigating the relationship between musical training and mathematical thinking in children. *Procedia − Social and Behavioural Sciences, 55*, 1134–1143.

Schellenberg, G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511–514.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*, 1047–1055.

See, B. H., & Kokotsaki, D. (2015). *Systematic review of the impact of arts education on the cognitive and non-cognitive outcomes.* London: Educational Endowment Foundation.

Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: Evidence for a sensitive period. *The Journal of Neuroscience, 33*(3), 1282–1290.

Styles, B., Clarkson, R., & Fowler, K. (2014). *Evaluation of Rhythm for Reading.* London: Education Endowment Foundation.

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *Neuroscientist, 16*(5), 565–577.