Evaluation of the impact of London Music Masters' Bridge Project on pupil's academic attainment

Tera Allas
Joshua Monahan

December 2014

Pro Bono Economics Reg.
Charity No. 1130567

London Music Masters
Charity No. 1119372

Foreword

Pro Bono Economics is delighted to introduce this report, the results of the results of work undertaken by Tera Allas and Joshua Monahan.

Pro Bono Economics was founded in 2009 with the aim of bringing the skills of economists into the third sector, working pro bono. Many charities could benefit from the expertise of economists, particularly in helping to understand measurement, impact and value. We think that by bringing together economists and charities we can not only benefit individual charities, but also publish analysis that can help the sector more broadly.

The team would like to thank the following people for their time and assistance in carrying out this project:

- Richard Davies, who wrote the project proposal with Pro Bono Economics and London Music Masters, carried out the bulk of the literature review, provided initial analysis of the National Pupil Database and outlined the overall approach

- Karen Hancock, who helped hypothesise and identify key factors that have been shown to influence children's academic performance and suggested a number of useful sources for the literature review

- Sumit Rahman, who helped think about the different ways in which descriptive statistics could be used to scope, refine and design the quantitative evaluation methodology

- Claire Samson, who initiated and shaped the project at London Music Masters, wrote the project proposal with Pro Bono Economics and Richard Davies, arranged access to the National Pupil Database and steered the project in its early phases

- Peter Sellen, Economic Adviser at the Department for Education, who provided very helpful comments on the methodology and analysis, in a personal capacity.

Introduction from London Music Masters

London Music Masters believes that music is of inherent cultural value and that it can have a lifelong impact on the social, educational and cultural development of children and communities. We are working towards a world where everyone can access extraordinary music.

We are passionate about providing world class learning opportunities for children in inner-city schools and are dedicated to training an ambitious and talented team of music graduates to facilitate this. We are also committed to enabling artists of international standing take extraordinary music from the concert hall into the community. We support the professional development of some of the world's most exciting young artists and give them the opportunity to share their talents with the widest possible audience.

The central purpose of our learning programme – of which Bridge Project is the largest component – is to make a difference to the lives of children in our schools by providing access to musical tuition and opportunities that are out of reach of many families in the UK. Over the past six years we have started to see remarkable outcomes for the children in terms of musical development, widening of social networks, increased resilience and an ability to learn new – and difficult – skills. There is tremendous scope for the transferability of these skills to enhance measurably the rest of the holistic development of our students.

As an organisation London Music Masters has a strong record of research driven practice from its early collaboration with Dr Ghislaine Kenyon and ongoing research projects with St Mary's University and the Royal College of Music. However, with this particular evaluation study, we wanted to explore something different. There is increasing evidence that music making produces additional, non-musical cognitive benefits for children as well as grown-ups. While this is not the primary purpose for Bridge Project, we wanted to know whether it was possible to quantify, through rigorous, data driven research, the impact Bridge has had to date on the participating children's academic performance.

One of the things we discovered through this process was that three years worth of data is not sufficient to confirm comments from teachers, parents and our own observations about academic impact. It was also important for us to grasp the very real challenge of finding causal links between actions (i.e. instrumental music lessons) and effects (i.e. academic improvement, particularly in numeracy and literacy) in an environment as complex as a primary school. It may be that this is not possible for school-based programmes like ours with data currently available and that we may need to rely on studies which have greater control over the research environment to lead the conversation on educational impact.

We are extremely grateful to Pro Bono Economics for helping us undertake this work and for carrying it out to such a high standard. We look forward to taking up the recommendations they have made and working with them in the future.

Rob Adediran
Executive Director, London Music Masters

Evaluation of the impact of London Music Masters' Bridge Project on pupils' academic attainment

Table of contents

Executive Summary

The purpose of the Bridge Project is to promote musical excellence and strive for greater diversity amongst performers and audiences. The programme provides young children with weekly musicianship sessions. The primary aims of Bridge are therefore musical. However, literature increasingly suggests that music making can also produce non-musical benefits.

The objective of this evaluation study was to test this hypothesis on the children who have participated in the Bridge Project. Specifically, the analysis aimed to establish whether children who have participated in Bridge achieved higher academic performance at Key Stage 1 than "would otherwise have been expected", i.e., as compared to an otherwise similar control group. The analysis found no statistically significant effect. There were some differences in the results between the two schools included in the analysis. However, these could have been due to lack of data on other factors that explain academic attainment.

This result should not be interpreted to suggest that there is no effect: it merely means that any effect was not detectable, given the data available for the analysis. Educational outcomes are determined by a very large number of factors at individual, cohort, school, neighbourhood and family levels. Many of these factors are not easily observable or data on them are not easily accessible. Given the small number of schools, cohorts and children who have so far participated in Bridge, and the large number of other things happening in these children's lives, it is perhaps not surprising that no effect was detected.

With time, there will be a larger number of cohorts to analyse. From a data point of view, it would benefit future analysis if the Bridge programme were run in a larger number of schools. For any further evaluations, it will be important that London Music Masters tracks the children that take part in Bridge and considers obtaining further data through bespoke surveys outside (but matched to) the National Pupil Database. As more Bridge children reach Key Stage 2, it will also be valuable to analyse whether an effect can be identified at this level, as the impacts are likely to take a number of years to come through.

1. Introduction

1.1 Overview of the Bridge Project

The purpose of the Bridge Project is to promote musical excellence and strive for greater diversity amongst performers and audiences. It aims to achieve this by providing weekly musicianship and musical instrument lessons for all children in a particular year group at a participating school. Bridge specifically targets schools where opportunities to be inspired by classical music early on might otherwise be limited. The schools are in some of the most culturally, socially and ethnically diverse parts of London.

The programme starts in Reception where the children (aged 4 and 5) take part in weekly musicianship sessions. At Key Stage 1 (KS1) all students (aged 5-7) begin to learn the violin with those showing most promise continuing on into Key Stage 2 (aged 9-11). The majority of the teaching takes place during the school day to ensure that everyone can participate. The students have up to 2 hours of lessons each week in small groups (between 2 and 15) and participate in a rich programme of performances and creative workshops.

Bridge Project has developed a teaching method, The Bridge Approach™, which embraces the best of Kodaly, Suzuki and contemporary teaching practice. In this one-to-many method, each group of children is taught by a specifically trained professional musician. The repertoire-led approach encourages children to make a confident sound and good progress from the start, whether singing or playing an instrument.

The Bridge programme is already highly successful in its primary aim: inspiring children to participate in classical music and increasing the diversity of the group of children who have access to first-class music tuition. For example, students are now winning places in the National Children's Orchestra and leading specialist music schools as well as auditioning for the Royal College of Music Junior Department.

As we discuss below, the design of the Bridge Project has implications for its evaluation. The fact that the schools have not been randomly selected could introduce a bias in the results relative to the general population of schools. Because Bridge is offered to all children until KS1, there are no natural comparators within the same schools. The methodology adopted for evaluation therefore needs to address these issues.

1.2 Objectives of the evaluation study

The objective of this study was to analyse whether children who have participated in the Bridge Project have benefited from the programme academically. Specifically, the analysis measured whether Bridge participants improved more quickly between entering school and Key Stage 1 than an otherwise similar control group.

It should be noted that the principal aims of the Bridge programme are musical and that it has been designed to maximise outcomes on this dimension. Any academic effects -- either positive or negative[1] – would therefore be an additional effect over and above of the primary mission of Bridge.

1.3 Approach to the evaluation study

The study looked at academic performance at Key Stage 1 in four categories: reading and writing, speaking and listening, mathematics and science. Given the importance of prior attainment (see section 3.3), performance was measured relative to the children's initial assessments in Reception.

The "treatment group" was made up of children who had participated in the Bridge Project from Reception to the end of Year 2 (Key Stage 1). This comprised three cohorts of children at each of two primary schools, which we will call School A and School B.

The "control group" was drawn from the National Pupil Database (NPD) in such a way that it comprised an anonymised random pupil from somewhere in London with the same observable characteristics as each child in the "treatment group".

The observable characteristics controlled for were based on a literature review of key factors that predict academic performance and variables available in the NPD. Further detail is provided below in sections 3.3 and 4.3.1.

A second analysis was also performed, looking at the academic performance of each cohort that participated in the Bridge Project and comparing it to the performance of the same year cohort in a number of schools with similar characteristics to School A and School B.

Necessarily, the analysis was limited by the availability of data, as well as time and resources available to perform the literature review. In this context, how the analysis should be interpreted is discussed further in section 4.6.

1.3.1 The cohorts in the treatment group

All of the cohorts which had completed Key Stage 1 by summer 2012 were included in the analysis. To date, a total of 3 cohorts at each of the two schools meet these criteria. Table 1 below summarises the information about the cohorts.

[1] It is plausible that a programme like the Bridge Project could in some cases have negative effects, if it displaced time spent on activities that would have been more beneficial for academic attainment.

Table 1: Bridge Project cohorts included in the analysis[2] (treatment group)

Cohort	School	Year started Bridge	Year reached Key Stage 1	Number of pupils
A1	School A	2007/08	2009/10	25
A2	School A	2008/09	2010/11	31
A3	School A	2009/10	2011/12	31
All	School A			**87**
B1	School B	2007/08	2009/10	28
B2	School B	2008/09	2010/11	23
B3	School B	2009/10	2011/12	26
All	School B			**77**

[2] i..e. excluding pupils who joined or left the school during KS1.

2. Analytical approach

2.1 Overall approach

The evaluation study followed the following stages: scoping, literature review, high level data analysis, detailed analysis design, pupil level analysis, school level analysis and interpretation.

The scoping phase identified the research questions London Music Masters wanted the study to address and produced a formal project proposal. This can be found at Annex A[3].

The literature review was used to provide context for and help design the quantitative analysis. Its results are summarised in section 3 and Annex B.

The high level data analysis produced descriptive statistics on the two schools and on the cohorts that had taken part in the Bridge Project. These were used to gain an understanding of the data available and the options for creating suitable comparisons and control groups.

The detailed design of the quantitative analysis was iterative and generated the methodology for measuring academic performance and for creating a control group for the Bridge cohorts in each school.

Two different analyses were then performed: one comparing individual pupils' performance (pupil level analysis) and another comparing the entire Bridge cohorts' performance (school level analysis).

The outputs from each phase are outlined in more detail in the sections below.

2.2 Data sources

Data for the literature review was obtained on the internet based on a number of search terms which were refined as more information was assimilated. The project team also interviewed Karen Hancock, previously the Chief Economist at the Department for Education, to identify further information sources.

The quantitative data, both for the Bridge Project participants and the control group, was drawn from the National Pupil Database (NPD).

[3] Note that due to changes in personnel at London Music Masters and Pro Bono Economics volunteers working on the project, the final shape of the analysis was slightly different from the original proposal.

2.2.1 Data on Bridge Project participants

It was not necessary to separately identify children who had participated in the Bridge Project. Instead, it was sufficient to identify each year group in each school which had participated in Bridge, as per Table 1 above. It could then be assumed, given the way that the project is delivered, that every child in each year group had participated in the Bridge Project. This is essentially the way the Bridge Project has been designed and implemented in these two schools.

Clearly, children who had joined a year group after Reception, or left before they reached the end of Year 2 (Key Stage 1), would not have benefited from the full Bridge programme. However, because the National Pupil Database (NPD) contains unique anonymised identifiers for each child across different years, it was possible to track each (anonymised) pupil and exclude any "late joiners" or "early leavers" from the treatment group.

2.2.2 The National Pupil Database (NPD)

The National Pupil Database (NPD) is one of the richest education datasets in the world.

There are a range of data sources in the NPD (such as the school census) which provide information about pupils' test and exam results and progression at each key stage. The NPD also includes data on pupil characteristics, such as their gender, ethnicity, first language, eligibility for free school meals and information about special educational needs.

The NPD covers practically all pupils in practically all schools and colleges in England[4]. This means that, for each year group, it contains data on more than 500,000 pupils. Each pupil has an anonymised reference number that enables longitudinal analysis. Some of the data goes back to 1995/96, and the richer datasets are available from 2006/07 onwards.

Subject to an approved application, the Department for Education makes extracts of the data available for the purpose of promoting the education or well-being of children in England. This could be through conducting research or analysis, producing statistics or providing information, advice or guidance.

For the purposes of this evaluation study, London Music Masters made a successful application to obtain a sub-set of the NPD data relevant for analysing the academic attainment of pupils in the cohorts listed in Table 1 above. The data used in this analysis is described in more detail in section 4.

[4] The NPD contains data for all state schools in England, non-maintained special schools, sixth form and further education colleges and, where available, independent schools.

3. Findings from the literature review

3.1 Scope of the literature review

The purpose of the literature review was to provide context for and help design a robust methodology for the quantitative analysis.

The four main aspects of the literature review were:

- Learning from other research studies that have used the NPD about the issues and approaches encountered and employed when working with the data
- Informing the quantitative evaluation methodology by looking at other studies that investigate the impact of specific school interventions
- Identifying the drivers of academic performance that would need to be controlled for in order to estimate the incremental impact of the Bridge Project alone
- Gaining an understanding of the theory of change and key hypotheses by summarising research findings on the impact of music training on pupils.

The details of the literature review are provided in Annex B. The main points on other studies' findings on the impact of music training and the drivers of academic attainment are summarised below.

3.2 Findings on the impact of music training

A large number of studies suggest a link between music education and various positive attributes for students, such as improved learning skills. The strongest quantitative evidence is found on the following effects:

- Musical training improves skills that are helpful for the development of speech, language and vocabulary, such as sound and pitch recognition and memory
- Musical training also involves skills and habits that are helpful for enhanced learning in general, such as concentration, planning and perseverance
- Playing and listening to music has been shown to reduce stress and anxiety; it could be hypothesised that this could also lead to enhanced performance in school
- The non-musical benefits appear to be discernible only after some years of training and are positively correlated with the amount of time in musical study.

3.3 Findings on drivers of academic attainment

It is generally accepted that pupils' academic attainment is determined by a very large number of factors, some of which are innate and others environmental. Moreover, these factors interact in complex ways.

However, most of the literature agrees that the best predictor of a student's academic attainment at any one point in time is his or her attainment in earlier life. Our interpretation is that this variable, *"prior attainment"*, can be seen as a proxy that incorporates valuable information about other factors, some of which are hard to observe or collect data on. It therefore forms the basic starting point for most academic studies on academic attainment.

Other variables which have been shown to correlate with academic attainment are somewhat constrained by data availability. However, at least the following are routinely considered to have an impact on a pupil's attainment:

- *School level factors:* quality of leadership, school resources, age groups taught (e.g., presence of sixth form), type of school (e.g., single sex or co-educational)
- *Cohort level factors:* quality of teaching, relationship and communication between teacher and pupil, peer group effects (e.g., prevalence of poverty in peer group)
- *Pupil level factors:* socio-economic background (including parents' level of education, the family's income level, poverty/deprivation), parental support, within-year age, gender, ethnicity, home and/or first language, special educational needs

Note: Free school meals (FSM) are often used as a proxy for socio-economic background, or specifically, lower family income.

As discussed below, our evaluation methodology can only control for those variables which are included in the NPD or otherwise easy to identify. Inevitably, this means that some important drivers of academic performance (such as school leadership or parental support) are not accounted for in this analysis. This makes it difficult to distinguish the precise effect of the Bridge Project from underlying variation.

4. Findings from the econometric analysis

For this study we studied pupils at School A and School B schools who completed Key Stage 1(KS1) in 2009/10, 2010/11 and 2011/12. These were all cohorts who had participated in the Bridge Project from Reception year onwards.

We removed pupils from the analysis if they had moved either to or from these schools between starting school and the end of KS1 on the basis that they would not have received the full "treatment"[5]. In total, we had 164 pupils in the treatment group.

4.1 Descriptive statistics for Bridge Project participants

In this section we set out some descriptive statistics on Bridge Project pupils to give some context. As a reference group we give the statistics for the average London school. London, rather than national, comparisons have been used throughout because of the difference in performance of London schools over the last decade or two.

Table 2: Descriptive statistics for Bridge Project participants

Percent of children who have the following characteristic:	**Average London School**	**School A**	**School B**
Female	49	53	53
Ethnicity other than White	56	87	82
Free School Meals	28	34	40
Special Educational Needs	26	9	22
First Language not English	43	38	53

Compared to London schools on average, School A and B are more ethnically diverse and have more pupils on free school meals, an indicator of lower incomes. School A has a slightly smaller proportion of pupils whose first language is not English than the London average.

[5] In the pupil level analysis, we also removed school movers from the control group, in order to avoid any bias. As it turned out, moving schools did not have a statistically significant effect on academic outcomes.

School B, on the contrary, has slightly more pupils whose first language is not English. Both schools have lower than average numbers of pupils with Special Educational Needs, particularly School A.

4.2 Metrics used to measure academic attainment

When pupils start school in Reception year, they are assessed[6] on a large range of factors that vary from ability to focus on activities to counting to 20. These are scored individually and summed to form an overall score. For this analysis, we standardised this overall score relative to all pupils in London.

At the end of Key Stage 1 pupils are assessed on their maths, science, reading and writing, and speaking and listening. These are assessed on whether they meet certain levels of attainment. We coded the levels as follows:

- 0 = working toward level 1,
- 1= Level 1,
- 2 = Level 2C,
- 3 = Level 2B,
- 4 = Level 2A,
- 5 = Level 3,
- 6 = Levels 4 and above.

We then took the average of these scores across the four assessments to give an overall score of academic attainment. Finally, we standardised this overall score relative to all pupils in London. While it would have been possible to map Early Years Foundation Stage assessments (from Reception) to the separate assessments for maths, science, reading and writing, and speaking and listening at KS1, and to conduct the analysis separately for each skill, we felt that the extra granularity was unlikely to add value[7].

To measure the effect of the Bridge Project we took the difference between these two standardised scores as our dependent variable. This measures the improvement in performance, as assessed by the school, over the time period relative to peers. To put it another way, it measures the movement of pupils within the standardised distribution of attainment between the beginning and end of Key Stage 1. We use this metric rather than the absolute level of performance at the end of Key Stage 1 in order to control for differences in school intake. This is consistent with the finding reported above in section 3.3

[6] Early Years Foundation Stage Profile

[7] This is a choice that could be revisited for any follow-on work. However, given all the other challenges with the data – and in particular small number of schools and pupils in the treatment group – it may not be worthwhile.

that *"prior attainment"* is a key explanatory variable that needs to be controlled for in educational evaluation studies.

4.3 Econometric methodology

We used two different matching methodologies to create control groups against which to compare the academic performance of Bridge Project pupils: pupil level and school level analyses. These are explained below. For both analyses, we used t-tests[8] to compare the improvement in attainment of Bridge project pupils and control groups.

4.3.1 Construction of the control group for the pupil level analysis

The first method was to find exact matches for each pupil, using matching variables that were available in the NPD and the literature review indicated as important predictors of academic performance. Exact matches were possible due to the discrete nature of the control variables and the sheer size of the NPD. The control variables used were as follows:

- Gender (Male, Female)
- Ethnicity (White, Black, Chinese, Asian, Mixed, Other)
- Free School Meal eligibility (Yes, No)
- First Language not English (Yes, No)
- Special Educational Needs (Statemented, School Action, School Action Plus, None)

The control group for the pupil-based matching was the same size as the treatment group[9].

4.3.2 Construction of the control group for the school level analysis

By using exact matches, the pupil-based method above provides a control group as closely aligned to the make-up of the Bridge Project cohorts as can be achieved with available observed data. However, as identified in the literature review, peer effects on academic attainment are significant and, with the concentration of factors known to be correlated with lower academic performance at School A and B, it is possible that peer effects have dampened the estimated effect of the Bridge Project. In plain English, we worried that pupils from social backgrounds associated with poor academic performance would do worse due to peer effects in schools where that was the norm rather than the exception. We

[8] Simple comparison-of-means with unequal variances and unequal samples.
[9] It would have been possible to have a larger control group by matching each treatment group pupil to a larger number of otherwise similar pupils. However, while there are some improvements in statistical power from increasing the size of the control, these diminish rapidly if the treatment group remains the same size. This additional analysis would not have changed the findings and was therefore not included.

therefore performed a second analysis using school-based matching, which was designed to control for peer effects in addition to the other variables listed above.

The second method created a control group from schools that were most similar to School A and B. School level statistics were calculated for the same variables as above. Envisage assigning all schools in London to a point in 11-dimensional[10] space based on the variables used. The ten schools that were the shortest distance from the mid-point[11] of School A and B were chosen as the basis of the control group. The control group was made up of the pupils at those schools who completed Key Stage 1 in 2009/10, 2010/11 and 2011/12[12]. The number of pupils in the control group for the school-based matching was 1014.

4.4 Results from the pupil level analysis

We found no statistically significant (positive or negative) effect of the Bridge Project when taking the treatment group as a whole, nor when broken down into sub-groups determined by the control variables described above. Note that the large standard deviations suggest there is substantial variation unaccounted for by our observed variables relative to the differences in means between the treatment and control groups.

Table 3: Results from the pupil level analysis

	Control			Treatment			t-stat	p
	mean	s.d.	n	mean	s.d.	n		
All	-0.02	0.77	164	-0.09	0.84	164	0.82	0.41

The way to interpret these results is broadly as follows. The small negative means for both the control and treatment group would indicate that both groups' performance deteriorated very slightly (and not statistically significantly) relative to all pupils in London (the group used to standardise the scores). However, the p-value indicates that any difference between the control and treatment groups has a more than a 40% chance of having been produced by random variation. In other words, any differences would not be considered statistically significant.

The full results are provided in Annex C, Table C.1.

[10] For each of the five variables, one less than the number of possibilities.
[11] The analysis was done for each of School A and School B separately but there was over 50% overlap in schools identified by this method. School B was the 6th closest school to School A and School A the 20th closest school to School B.
[12] Excluding those who joined the school late, to ensure comparability with the treatment group.

4.5 Results from the school level analysis

Similarly, repeating the analysis using the school-based matching, there were no statistically significant results.

Table 4: Results from the school level analysis

	Control			Treatment			t-stat	p
	mean	s.d.	n	mean	s.d.	n		
All	-0.03	0.74	1014	-0.02	0.78	164	0.16	0.87

Again, both the control and treatment group deteriorated very slightly relative to the wider London population. The difference between the control group and the treatment group is so small that there is an 87% chance it was produced by random variation.

Full results are in Annex C, Table C.2.

4.6 Interpretation of the results

Our econometric analysis did not provide quantitative evidence that the Bridge Project was having either a negative or positive effect on the academic attainment of pupils.

However, absence of evidence is not evidence of absence of an effect. The small sample size and large degree of variation unaccounted for by observed variables could easily disguise even moderate effects.

Furthermore, the nil result reported here is, to an extent, a product of chance. If we analyse the attainment of pupils at School A and B separately, we find that School A pupils did better (mean = 0.34) and School B pupils did worse (mean = -0.58) than the control group (mean = -0.02). Both differences are statistically significant (p-values of 0.0003 and 0.0000007, respectively). In aggregate, these cancel each other out, resulting in the reported nil result. Rather than a genuine result, this could have been due to a chance: analysing one school doing particularly well and one doing less well at a particular point in time.

We would caution against reading too much into analytical results from small numbers of schools when we know from our literature review that school level effects matter. We would need to analyse sufficient numbers of schools for these effects to “wash out” to be able to say something more robust about the effect on academic outcomes of a programme like the Bridge Project. Alternatively, we would need to be able to control in other ways for some of the key school level characteristics, such as leadership and teacher quality.

The small sample size, the small number of schools, the differing results for the two schools and the large degree of unaccounted for variability make it difficult to make robust statements. In the end, statistical analysis is a fairly weak tool in determining the effect of projects like Bridge on academic outcomes at the small scale on which they operate.

5. Next steps

Despite the non-conclusive results from this evaluation study, we believe that there is value in considering further ways to measure the effects of the Bridge project on pupils' academic performance. This belief is based on the relatively strong, and growing, academic literature that suggests that such effects are real and positive. It may also be valuable to perform an additional literature review specifically focused on the types of musical training, environments or student populations that tend to yield the most positive results.

In terms of the quantitative analysis, there are several ways in which this could be strengthened over time and which might produce more concrete results. A number of these further steps are outlined in the original project proposal in Annex A but it has not been possible to implement them as part of this evaluation study. Based on our experience to date, we would suggest that at least the following are worth further consideration:

- *More data:* If it were possible to gather data on some of the key missing explanatory variables which are known to explain variability in academic attainment[13], this would likely result in clearer results on the incremental impact of Bridge alone. In addition, it is possible that there are more accurate, less subjective, ways of assessing pupils' academic performance in the Early Years and Key Stage 1 than those contained in the NPD. However, capturing additional data for Bridge pupils, and a control group, would likely incur costs (either in money or time) and might run into data protection issues[14]. We would suggest a further, more thorough, literature review in order to assess the likely costs and benefits of any particular additional data.

- *More schools, more cohorts and more pupils:* The more schools Bridge Project is run at, the more it can be expected that the school level factors that are difficult to control for (such as leadership quality) will "wash out". However, our judgement is that the number of schools that would be required for this statistical effect to be achieved is much larger than the number of schools where Bridge is realistically likely to be run in the near future. Even with a small number of schools, though, as the sample size of pupils and cohorts who have taken part in Bridge increases, any effects on academic performance might become more detectable.

- *More time:* The literature review indicates that the non-musical benefits of music training can take several years to become apparent. It is therefore likely to be

[13] Such as within-year age, parents' educational background, parental support, family income, neighbourhood characteristics (e.g., levels of unemployment, average level of education) and quality of school leadership.
[14] Any data captured outside the National Pupil Database would need to be matched with the anonymised data in the NPD. This is possible using the pupil reference identifiers in the NPD but would involve more steps, and hence time and possibility of errors, than using purely NPD data. Some of the data one would ideally control for in the analysis is of sensitive nature, so handling it would require expertise in data protection issues.

worthwhile to analyse the impact of the Bridge Project on pupils' academic attainment at Key Stage 2. It should be noted, however, that because only selected pupils continue as part of the Bridge Project after Key Stage 1, the sample size is likely to be even smaller in this case. On the other hand, it would be possible to construct a control group from within the same schools. This would introduce additional issues (e.g., smaller universe from which to find matches for the Bridge pupils) but should in theory control for some school and cohort level factors.

- *Other approaches:* Given the inherent difficulty of detecting a statistical effect for a relatively small and non-selective programme such as Bridge, London Music Masters should not discount other approaches. For example, while the results could only be considered indicative, it is worthwhile continuing to gather – systematically and ad hoc – both teachers' and parents' feedback on the impact they believe the Bridge Project is having on pupils' academic performance. Were LMM to want to conduct slightly more structured surveys on this, it would be well worth engaging social researchers or other experts to formulate the questions in ways that are likely to be most illuminating (and result in least biased answers).

At a minimum, we would recommend that London Music Masters find a way to track and keep a database of Bridge Project pupils over time for future evaluation purposes[15]. Information that is likely to be helpful in the future includes: date when the pupil started and finished Bridge, number of hours of tuition per week in each year the pupil was part of Bridge, and reference information allowing the pupil to be matched with data in the NPD. As indicated above, LMM may then decide to complement this information with other information gathered via, for example, bespoke teacher, parent or pupil surveys.

[15] Note that this may require expert advice on data protection issues.

ANNEXES

ANNEX A: ORIGINAL PROJECT PROPOSAL

Richard Davies and Claire Samson

A.1 Overview

The aim of this study is to measure the extent of the **non-musical benefits**[16] attributable to London Music Masters (LMM)'s Bridge Project (BP).

There have been a number of analyses of projects that have similar aims to the Bridge Project.[17] The differentiating factor with this study is that it will be explicitly quantitative. While LMM will continue to internally conduct case studies of particular children and their 'musical stories' (as other studies have relied heavily upon), this will not be the focus of the study undertaken by PBE. Rather, we will conduct a rigorous quantitative impact test, identifying a suitable control group and explicitly controlling for various biases that could invalidate our results.

The project should be useful for LMM, as part of its own evaluation; it should also have wider benefits for other charities engaged in related work, none of whom has yet – to our knowledge – carried out a statistically rigorous piece of work.

A.2 High level approach

We suggest the study would have three main strands:

- **Strand 1**: To obtain and analyse whatever data is available covering the first 4 years of the BP (2008-2012). To identify a comparator group of children—the control group—using the National Pupil Database. Using this back-data to set out some stylised facts and conduct some initial tests of the Bridge Project's non-musical impact to date.
- **Strand 2**: To identify the ideal set of data we need in the near term, should there be any limitations to the output of stage 1. This may, for example, include using a questionnaire, or agreeing to access slightly finer and more pupil specific data so that we can track children over time. If any changes to data collection are needed, we would aim to put new collection in place by September 2012, to allow analysis at the end of the 2012/13 school year.

[16] LMM is confident in its ability to assess the **musical** benefits of the programme, so this will not be included in the scope of this study with PBE. However, the PBE representative may offer advice to help LMM analyse the musical benefits through appropriate statistical methods.
[17] For example, the In Harmony projects.

- **Strand 3:** To consider data collection requirements over the longer term – i.e. appropriate markers to assess the ongoing non-musical benefits of the BP as its students progress to secondary school, university and/or the workplace. To draw up a timeline to indicate – over the course of a child's involvement with the BP and beyond – what data should ideally be collected, how, and when. Ensuring that the LMM team are aware of any methods by which data has been collected and analysed during the project with PBE such that they can continue to do this independently in the future.

A.3 Data needed and likely analysis

At our initial meeting we (LMM and PBE) agreed about the need for rigorous analysis. The majority of the work involved in the study will be establishing exactly which data we need, how to access it, and then how to analyse the data. This section of the proposal sets out our initial thoughts in this area.

Literature review and hypotheses to test

We will need to conduct a targeted literature review in order to examine the methodologies used in other studies. This should include two elements:

(1) A summary of other peer reviews.
(2) Studies of other interventions in schools to establish the generally accepted group of variables used in performance studies in schools.

Using these reviews we will be able to identify more precise hypotheses to test. One idea, for example, is that musicianship training helps to improve concentration and that concentration is crucial for writing a coherent paragraph, concentrate on a book, or complete a maths problem. We need to ground these intuitive ideas in the academic literature. This will give us a theory for why music should lead to better outcomes in other subjects. We will then test this theory using data.

Measuring non-musical impact: academic benefits

LMM has already conducted some preliminary analysis of the potential academic benefits of partaking in the BP – namely, looking at the percentage of BP students achieving the nationally expected grade or higher at end of Key Stage 1 (Year 2), compared to a control group, in the following disciplines:

- Reading
- Writing
- Maths

We propose that we limit our study of academic benefits to these variables.[18]

Measuring non-musical impact: social/behavioural benefits

The social/behavioural benefits of music education are compelling to funders and are often cited in a qualitative way by other organisations, through quotes and case studies. LMM is keen to find a way of quantifying the analysis of these benefits to assess them more rigorously. This study will aim to produce questionnaires for children/parents/teachers which will allow LMM to gather data on non-musical performance variables and more intangible variables – for example confidence, concentration, happiness, anxiety, measures of behaviour. The study will aim to set out a methodology through which this data may be quantified and analysed to give meaningful insights rather than simply anecdotes.

The treatment group

The treatment population is the group of children that are currently, have been, or will be involved with the Bridge Project. It consists of children attending LMM's three partner primary schools.

The control group

One of the most important things for the study is to identify an appropriate control group(s). We need a group of children that are as similar as possible to the children in the treatment groups. Ideally, however, they would be from different schools (i.e. not one of the Bridge Project schools) since there might be spill over benefits for even non Bridge Project children from having the Bridge Project involved in their school.

We propose to use the National Pupil Database (NPD) to identify suitable control groups. That is we would match a number of schools based on objective quantifiable characteristics. This could include school and class size, the relative prosperity of the area and the number of children that quality for free school meals, for example.

We may also seek to identify suitable matches based on conversations with teachers in the sample schools, to ascertain which other schools in the locality are similar. These subjective suggestions would be cross-checked using the NPD as above.

[18] Narrowing our choice of dependent variables to these will not lead to a narrow study. Even if we were to choose just one variable—performance in maths say—the analysis will need to be nuanced. For example, suppose we are interested solely in music and maths. We would then need to decide whether it is the absolute maths performance level (% of correct answers) or whether it is performance against a school benchmark (% correct, relative to school target) or whether performance against a personal benchmark (% correct, relative to predicted grade) that is most relevant. Probably we would look at all three of these. We would then need to control for various effects using the control group, and the 'cleaned' data in our final tests. This means one performance variable produces a lot of different things to look at.

Historical data analysis

We will need to collect two sets of data. First we need to collect back data that are held by the National Pupil Database. LMM has attempted to access this data before but the request was rejected. We will re-request data from the NPD in the early stages of this study, having sought advice on the key criteria for a successful application.

Forward looking data collection

Second, we want to put in place new data collection, targeted for the purposes of future BP impact studies. A literature review will help clarify the ideal set of data that is needed. This will help us design a template and a timeline for any new data that we want to collect. We will also aim to ascertain a suitable timetable for re-sampling of control groups, again based on examination of relevant literature.

Ensuring rigorous results

We are aware of the shortfalls in other analyses conducted in this area. Simply asking if there is a correlation between music lessons and performance is problematic for various reasons. The main problem with this analysis will be in "identification". That is, we need to be sure that we can correctly disentangle differences in performance that are due to the BP and those that are not. Some common pitfalls the study will seek to avoid are set out below.

1. **Sample selection bias**: This would occur if the schools that take part in the Bridge Project are somehow systematically different to other schools: they might have more ambitious parents or teachers, for example.

2. **Omitted variables bias**: We need to make sure that no explanatory factors that might mean our control and treatment groups perform differently are omitted from our analysis. We can avoid this by using a rich data set, with as much relevant information from the NPD collected as possible. The study should also cross check what is done in other educational attainment studies.

3. **Endogeneity bias**: This would occur if attainment levels determine the extent to which BP is involved in a school. This would mean it would be difficult to disentangle the effect that the BP has on standards from the effect that standards have on attracting the BP. To clarify this we should include some information on the process that the BP goes through in identifying partner schools and appropriate control groups.

A.4 Stages and timing

Section stage	*To do*	*Dates*
Literature review	• Review educational literature • Speak to relevant academics	• Aug 2012 • Sep/Oct 2012
Hypotheses to test	• Set out hypotheses • Agree methodology	• Aug 2012 • Aug 2012
Backward looking data analysis	• Prepare application for National Pupil Database • Receive data and undertake statistical analysis	• Summer-Autumn 2012 • Spring-Summer 2013
Setting up forward-looking data collection:		
a) Social & behavioural aspects	• Investigate pre-existing tests / questionnaires to assess social & behavioural changes; seek advice from relevant academics • Prepare draft questionnaires and seek advice from head teachers & parents • Pilot with selection of parents • Refine & deliver final versions of questionnaires to parents • Construct timetable for future questionnaire distribution & analysis •	• Aug-Sep 2012 • Autumn 2012 • Summer 2013 • Sep 2013 • Autumn 2013
b) Academic results (quantitative)	• Define strategy & construct timetable for information requests in order to undertake annual analysis at appropriate points (KS1, KS2, GCSE, A-level etc.)	• Summer 2013
Initial write-up	• Write up report	• Autumn 2013

ANNEX B: LITERATURE REVIEW

Richard Davies and Tera Allas

This annex summarises academic literature relevant for the investigation of the impact of the Bridge Project on pupils' academic attainment. It focuses on four types of study:

1. Studies that use the National Pupil Database (NPD), the principal dataset that we have used our study. These help refine the most effective approach to analysing the data.

2. Empirical studies that investigate the impact of interventions—i.e. some change in policy—in schools. We are interested in any change—for example a change in curriculum, or a new approach to physical education—and how researchers estimate the impact of the change, especially in studies that use a clearly specified control group and robust empirical methodology.

3. Research that identifies variables that correlate with children's academic attainment. These drivers could be at national, local, school, cohort or individual level. Identifying these drivers is important for ensuring they are controlled for in our analysis. Because of national differences, we have mainly focused on UK studies.

4. Research that looks (using any empirical technique) at the impact of music training on pupils' skills and attainment. These help understand the underlying theory of change and develop clear hypotheses to test in the econometric analysis.

B.1 Studies using the National Pupil Database (NPD)

The NPD includes information, at pupil level, on pupil and school characteristics and on pupil attainment for practically all school children in England. It is a rich dataset, and has the important property that it is longitudinal, so that using it Bridge Project students can be tracked year on year. It is also based on the UK census so that it measures the entire population (rather than just a sub-sample) of students. The database is widely viewed by the research community as high-quality. For example, researchers point out that over 98% of pupils have valid ethnicity records, and over 99% have accurate data on whether English is the pupil's first language or not. The properties of the dataset, and its wide use in policy research (including by the Department for Education) means we can be confident in basing our study largely on this data.

The NPD features in a number of relevant pieces of research. Simpson, Jivraj and Marquis (2011) look at how the NPD can be used to track pupils as they move around the country. The focus of their study is internal (and international) migration. The study uses the NPD variable on free school meals as a proxy for poverty; it finds that students from families in poverty and more likely to move than others. The paper also describes how the Pupil

Matching Reference can be used to track children over time, to create a longitudinal dataset.

Bell et al (2009), looks at the exam results of pupils following different science exams (GCSE Applied Science and Double Award Science) on the attainment between Key Stage 3 and Key Stage 4. The empirical analysis used NPD data for pupils in a large number (460) of secondary schools where both types of science exam were used. The students in the study were taking GCSE sciences in 2003-2005; these were the second cohort taking the Applied Science curriculum, which was new at that point. The findings are quite specific to science, and a number are not very relevant for our study. One point from this (and other) studies is that it is important to control for the pupil's sex; in this study girls did better in Applied Science than their KS3 results would have predicted. The interesting point is that the NPD can clearly be used to identify different sets of pupils and that KS results can be used as predictors of future results

B.2 Empirical tests of intervention in schools

The Department for Education (2012) recently published a report that looked at the impact of the "Sure Start" program. This is a good comparator study for ours because of its control group approach: the study it investigates child "functioning" in 150 Sure Start areas, and compares these to families in those disadvantaged areas that do not have a Sure Start project. The paper has longitudinal strategy: previous studies had assessed children at 9 months, 3 and 5 years old. The 7-year-old study then randomly selected a subset of the children and families that had been previously studied. Interestingly, the study of 7 year olds supports the hypothesis that supporting children and families improved parent outcomes but there was no evidence of better child outcomes. Another interesting finding for our project is that benefits were persistent. The effects on parents lasted at least two years after their last contact with Sure Start programmes. According to the authors, there are many examples where this does not happen, and after the short-term benefits of a social intervention, the impact can decay or "wash-out".

The Department for Education (2010) report on the evaluation of the UK Resilience Programme provides another good example of an empirical study examining the effects of a specific intervention. The aim of the study was to investigate whether the Resilience Programme had an impact on children's well-being, behaviour, attendance and academic attainment. While the information on pupils' well-being was collected through pupil questionnaires, data on academic attainment was based on returns from schools and the data on absences was based on the NPD. What is interesting about the findings of this study is that, while there is a significant observed positive effect of the programme on pupils' attainment in English, this falls to zero once pupil characteristics and the school baseline are controlled for. Another interesting feature is that the results appear to fade over time.

These findings reinforce the importance of using a carefully selected control group to analyse effects; and of looking at effects over time.

B.3 Research on drivers of academic attainment

There are two types of studies that are helpful in understanding all the variables that might have an impact on pupils' academic attainment and that need to be controlled for in evaluating the specific impact of the Bridge Programme. Firstly, some studies are directly aimed at understanding differentiating factors that explain pupils' attainment. Secondly, it is useful to look at which variables other evaluation studies have controlled for. Up to a point, however, it has to be recognised that both types of studies have often been constrained by what data is available or attainable.

Jenkins (2009) provides a comprehensive quantitative analysis of factors affecting attainment at GCSE. A good overview is also provided in Allen (2013), even though this study focuses on how to close the gap between socio-economically disadvantaged and other pupils. Slater (2009) analyses specifically the impact of teacher effectiveness, controlling for other variables. These and other studies identify (within the context of their research methodology) the following variables as having statistically significant impacts on academic attainment:

- *Prior attainment* (see below)
- *School level factors:* quality of leadership, school resources, age groups taught (e.g., presence of sixth form), single sex vs. mixed, grammar vs. comprehensive, denominational vs. non-denominational
- *Cohort level factors:* quality of teaching, relationship and communication between teacher and pupil, peer group effects (e.g., composition of cohort in terms of deprivation or ethnic diversity)
- *Pupil level factors:*
    - *Neighbourhood* where the pupil lives: unemployment, low skills population, lone-parent households
    - *Family* of the pupil: socio-economic background (including poverty), parents' level of education, parental support
    - *Individual attributes:* within-year age, gender, ethnicity, home and/or first language, special educational needs

It is worth specifically commenting on the importance of controlling for *prior attainment*. Most studies either find or take as given that one of the most important predictors of a student's academic attainment this year is his or her attainment in earlier life. Our interpretation is that this variable can essentially be seen as a proxy that incorporates valuable information about other factors, some of which are hard to observe or collect data on. It therefore forms the basic starting point for most academic studies.

It is also worth noting that evidence (Department for Children, Schools and Families 2009) shows that deprivation in particular affects both children's starting point attainment as well as progress over time – in other words, an initial gap in attainment gets wider over time. This may partly explain why both of the schools in this study "deteriorated" relative to other schools in the time period analysed.

B.4 Studies looking at the impact of music training

Annex 3 of the 2011 Department for Education report "Importance of Music" includes a summary of the findings of studies looking at the impact of music training on pupils' academic and social development. The summary draws heavily on Hallam (2010), which provides a comprehensive survey of the literature. A number of points stand out:

a. There is a strong link between musical skills and speech. Analysis of very young (pre-school) children found manipulation of speech sounds and reading were better in children with musical training; also in reading, and phonological awareness.

b. The process seems to be "transfer" of skills. For example Gromko (2005) looked at nursery age children who took music lessons that included emphasis on beat, rhythm and pitch and associated the sounds with symbols. These children had better phonemic awareness than those in the control group.

c. The benefits also apply to language. 8 year olds with musical training beat non-musicians on both music (expected) and language (potentially an example of indirect transfer). This seems to be because the brain's development of pitch processing is occurs earlier in music than in language.

d. Vocabulary seems to improve with music lessons too. Another study looked at the relationship between piano lessons and vocabulary. This study used a control group in a similar way to our plan. 46 children were in the treatment group. They studied piano for 3 consecutive years. 57 were in the control group, and did not have piano training. The pianists ended up with better vocabulary and verbal sequencing. That said, the effects had not manifested themselves after the first two years. One explanation could be that relatively extended period of music instruction is needed. This is another reason that the NPD, which allows a longitudinal study, is a good option for us.

e. The links between music and mathematics are less clear. Some studies Hallam reviews have found no impact at all. Other studies do find a relationship. On has interesting parallels with the Bridge Project. Catterall et al. (1999) used American data to examine poor children that were strong in maths at 12th grade. They found that many [a third] of these were taking music lessons.

f. A common theme among studies looking at the academic impacts of music tuition is that the benefits appear to be the strongest for children whose starting-point skills are the lowest. Greatest improvements were seen when teaching was tailored to pupils' existing skills and abilities.

g. Where studies have looked at it, they have found that rhythmic music training appears to be the most effective in improving reading and maths skills and IQ scores.

There are a number of other interesting findings outside Hallam's review.

a. Kraus and Chandrasekaran (2010) show that music training develops skills that are important in non-musical tasks. Examples of strengths that the musically trained possess include: extra-linguistic speech functions (things like being able to tell the emotion of a speaker, and whether a speaker has just made a statement or asked a question) and a better ability with foreign languages where it is important to quickly be able to pick up the importance of differences in tone specific to that language. The authors note that in all of the studies benefits were correlated with the amount of time in musical study, again justifying using the longitudinal nature of the NPD.

b. Schellenberg (2004) investigates the link between music lessons and the Intelligence Quotient (IQ) test. This paper, like our project, is explicitly looking for indirect or "collateral" benefits that "extend to non-musical areas of cognition". These, again, are called "transfer" effects. Schellenberg points out that music lessons and playing requires extended periods of focus, attention, daily training practice, reading music [a new language] and memorizing lengthy passages, and fine-motor skills. The hypothesis is that all this spills over or transfers to other areas of cognition. Moreover, children's brains are highly "plastic" and sensitive to environmental changes. The general problem, Schellenberg points out, is that children with active and ambitious parents are both more likely to do better in IQ tests and to take music lessons. So in the music/IQ testing literature there is going to be a problem of spurious correlation. So to test the hypothesis, the study separated (randomly) 144 children into four different groups, some receiving musical (keyboard or voice) training, and two different control groups, one that received drama lessons, and one that received no extra lessons. The musicians had IQ increases of 7 points, while the control group had IQ increases of 4.3 points.

c. Walker and Boyce-Tillman (2002) look at the impact music lessons may have for children suffering from chronic anxiety. The paper is a series of five case studies. The children all had severe chronic anxiety disorders. The children, their parents, music-teachers and

family therapists all recorded observations. The results are interesting. In just one year, music lessons were associated with better thoughts and feelings (self-confidence, social ease, independence, control over intrusive thoughts and feelings). This methodology is very involved and qualitative however, and so probably beyond the scope of our project. Instead, one of the tests that Schellenberg uses with parents and teachers (the "BASC" that measures "maladaptive and adaptive social functioning" including six measures (Hyperactivity, Aggression, Anxiety, Depression, Atypicality, and Attention Problems) are combined to form a "Behavioural Symptoms Index". We could investigate using this, and whether the children's teachers (or their parents) would be willing to fill this out.

d. There is a literature on the relationship between sport and brain function. Studies from this literature are useful for our study as there are strong parallels with our hypothesis. The main parallel is that the treatment/benefit relationship is indirect. That is, the treatment (sport) gives a direct benefit (fitness) and, it is argued, an indirect one (better concentration, school performance etc.). Hillman et al (2004) is an example of the sport/concentration literature. The paper tests exercise and "executive control" (a specific type of brain function and processing) in older people. The authors use neuroscience-based methods, looking at electroencephalography to measure differences in cognitive function across age and physical activity levels. The researchers examine brain activation time as a measure of brain activity in response to or preparation for a stimulus or response. Clearly, something like this is beyond the scope of our study, but it is interesting to note that there is relatively "hard" science that would support the notion that interventions like sport and music can have indirect beneficial impact on performance.

   Some of the sport and education studies should be treated with caution, however. There is a large industry in providing (at a cost to the public sector) services such as "Brain Gym". Hyatt (2007) provides a sceptical analysis of whether Brain Gym actually works. Others are more explicit in noting that support for this type of intervention in schools is based on bogus studies; Ben Goldacre has written about this in The Guardian recently.

B.5 References

Allen, T. (2013). "Research paper for Ofsted's 'Access and achievement in education 2013 review'". Institute of Education, University of London. http://dera.ioe.ac.uk/18562/1/Access%20and%20achievement%20-%20Drivers%20and%20barriers%20to%20raising%20achievement.pdf

Anvari S.H., Trainor L.J., Woodside J. & Levy B.Z. (2002). "Relations among musical skills, phonological processing, and early reading ability in preschool children". Journal of Experimental Child Psychology, 83, 111-130.

Barnett, S.M. & Ceci, S.J. (2002). “When and where do we apply what we learn? A taxonomy for transfer”. Psychological Bulletin, 128, 612-637.

Bell, J., Donnelly, J., Homer, M. & Pell, G. (2009). “A Value-Added Study of the Impact of Science Curriculum Reform Using the National Pupil Database”. British Educational Research Journal, volume 35, number 1, February 2009 pages 119-135.

Burns, M.K., (2008). “Research-Based Academic Interventions”.

Catterall, J., Chapleau, R. & Iwanga, J. (1999). “Involvement in the arts and human development: General involvement and Intensive Involvement in music and theatre arts”. Champions of Change: The impact of the Arts on Learning, Washington DC Arts Education Partnership.

Chernick, A (2009). “The Effects of Movement Based Intervention Programs on Learning in Grades K-12”.
https://www.nmu.edu/sites/DrupalEducation/files/UserFiles/Files/Pre-Drupal/SiteSections/Students/GradPapers/Projects/Chernick_Alycia_MP.pdf

Department for Children, Schools and Families (2009). “Deprivation and Education: The evidence on pupils in England, Foundation Stage to Key Stage 4”
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/222172/DCSF-RTP-09-01.pdf

Department for Education (2010). “UK Resilience Programme Evaluation: Final Report”. Research Report DFE-RR097.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182419/DFE-RR097.pdf

Department for Education (2011). “Importance of Music – A National Plan for Music Education”.
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/180973/DFE-00086-2011.pdf

Department for Education (2012). “The impact of Sure Start Local Programmes on seven year olds and their families”. Research Brief 220.
https://www.education.gov.uk/publications/eOrderingDownload/DFE-RB220.pdf

Ewens, D. (2005). “The National and London Pupil Datasets: An introductory briefing for researchers and research users”. Greater London Authority, DMAG Briefing 2005/8.

Flora, C. (2005). “Fit 'N' Brainy”. Psychology Today, 98, 121.

Gromko, J. (2005). “The effect of music instruction on phonemic awareness in beginning readers”. Journal of College Reading and Learning, 53(3), 199-209.

Hallam, S. (2010). “The power of music: its impact of the intellectual, personal and social development of children and young people”. International Journal of Music Education 38(3), 269-289.
http://www.devon.gov.uk/dms-powerofmusic.pdf

Harvard, Project Zero. Retrieved from http://www.pz.harvard.edu/Research/MISchool.htm

Hillman, C.H., Belopolsky, A. V., Snook, E. M., Kramer, A. F. & McAuley, E. (2004). “Physical Activity and Executive Control: Implications for Increased Cognitive Health During Older Adulthood”. Research Quarterly for Exercise and Sport, 75(2), 176-185.
http://kch.illinois.edu/research/labs/neurocognitivekinesiology/files/Articles/Hillman_2004_PhysicalActivityAndExecutive.pdf

Hyatt, K. (2007). “Brain gym: Building stronger brains or wishful thinking?” Remedial and Special Education, 28(2), 117-24.

Jenkins, A., Levacic, R & Vignoles, A. (2006). “Estimating the Relationship between School Resources and Pupil Attainment at GCSE”. Research Report RR727, Department for Education and Skills.
http://eprints.ioe.ac.uk/1775/1/Jenkins2006Estimating.pdf

Kounali, D., Goldstein, H., Robinson, A. & Lauder, H. (2007). “The Probity of Free School Meals as a Proxy Measure for Disadvantage”.
www.cmm.bristol.ac.uk/publications/FSM.pdf

Krau, N. & Chandrasekaran, B. (2010). “Music training for the development of auditory skills”. Nature Reviews Neuroscience 11, 599-605 (August 2010).
http://www.soc.northwestern.edu/brainvolts/documents/KrausChandrasekeran_NRN10.pdf

Love, J., Kisker, E. E., Ross, C. M. Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni & A. S., Brady-Smith, C. (2002). “Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start”. Volume 1: Final Technical Report. Princeton, NJ: Mathematica Policy Research Inc.
http://www.mathematica-mpr.com/PDFs/ehsfinalvol1.pdf

Magne, C., Schon, D. & Besson, M. (2006). “Musician children detect pitch violations in both music and language better than nonmusician children: behavioural and electrophysiological approaches”. Journal of Cognitive Neuroscience, 18: 199-211.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. & Malin, A. (2008). “Preschool influences on mathematics achievement”. Science, 321, 1161-1162.

Peynircioglu, Z., Durgunoglu, A.Y. & Uney-Kusefoglu, B. (2002). “Phonological awareness and musical aptitude”. Journal of Research in Reading, 25(1), 68-80.

Piro, J.M. & Ortiz, C. (2009). “The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students”. Psychology of Music, 37(3), 325-347.

Plewis, I. & Kallis, C. (2008). “Changing economic circumstances in childhood and their effects on subsequent educational and other outcomes”. Department for Work and Pensions WP 49.

Schellenberg EG (2004). “Music lessons enhance IQ”. Psychol Sci. 2004 Aug;15(8):511-4. http://www.ncbi.nlm.nih.gov/pubmed/15270994

Simpson, L., Jivraj. S., & Marquis, N. (2011). “Measuring Internal and International Migration from the National Pupil Database”. Cathie Marsh Centre for Census and Survey Research, University of Manchester.
http://www.communities.gov.uk/documents/corporate/pdf/1854829.pdf

Slater, H., Davies, N. & Burgess, S. (2009) “Do teachers matter? Measuring the variation in teacher effectiveness in England”. Working Paper No.09/212, Centre for Market and Public Organisation, Bristol Institute of Public Affairs.
http://www.bristol.ac.uk/cmpo/publications/papers/2009/wp212.pdf

Strand, S. & Demie, F. (2007). “Pupil mobility, attainment and progress in secondary school”. Educational Studies, 33(3), pp.313-331.

Vandell, D.L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD Early Child Care Research Network (2010). “Do Effects of Early Child Care Extend to Age 15 Years? Results from the NICHD Study of Early Child Care and Youth Development”. Child Development, 81, 737-756.

Walker, J. & Boyce-Tillman, J. (2002). “Music lessons on prescription? The impact of music lessons for children with chronic anxiety problems”. Health Education, Vol. 102 Issue: 4, pp.172 – 179.
http://www.emeraldinsight.com/journals.htm?articleid=871808

ANNEX C: STATISTICAL RESULTS

Table C.1: Pupil-level analysis

	Control			Treatment			t-stat	p
	mean	s.d.	n	mean	s.d.	n		
All								
All	-0.02	0.77	164	-0.09	0.84	164	0.82	0.41
Gender								
Female	-0.09	0.76	87	-0.17	0.84	87	0.61	0.54
Male	0.06	0.78	77	-0.01	0.84	77	0.55	0.59
Ethnicity								
Black	-0.05	0.73	97	-0.07	0.85	97	0.18	0.86
Mixed	0.08	0.91	28	-0.15	0.95	28	0.91	0.37
White	-0.13	0.72	25	-0.12	0.70	25	0.03	0.98
Asian	0.18	1.00	5	-0.18	0.57	5	0.71	0.52
Chinese	-0.32	0.28	2	0.12	1.24	2	0.48	0.71
Other	0.34	0.91	7	-0.06	1.05	7	0.75	0.48
Free School Meals								
Yes	-0.015	0.84	61	-0.23	0.88	61	1.38	0.17
No	-0.02	0.73	103	-0.01	0.81	103	0.11	0.91
First Language								
English	-0.04	0.80	90	-0.19	0.85	90	1.22	0.22
Other	0.00	0.75	74	0.02	0.82	74	0.16	0.87
Special Educational Needs								
None	-0.04	0.76	139	-0.07	0.83	139	0.26	0.80
School Action	0.32	1.13	9	-0.57	0.99	9	1.78	0.11
School Action Plus	-0.14	0.53	14	-0.13	0.81	14	0.04	0.97
Statemented	0.78	0.95	2	0.47	0.23	2	0.46	0.73

Table C.2: School-level analysis:

	Control			Treatment			t-stat	p
	mean	s.d.	n	mean	s.d.	n		
All								
All	-0.03	0.74	1014	-0.02	0.78	164	0.16	0.87
Gender								
Female	-0.05	0.71	512	-0.10	0.77	87	0.46	0.64
Male	-0.01	0.76	502	0.06	0.78	77	0.74	0.46
Ethnicity								
Black	-0.06	0.71	546	0.00	0.78	97	0.77	0.44
Mixed	-0.05	0.71	139	-0.08	0.89	28	0.17	0.87
White	-0.03	0.81	230	-0.06	0.64	25	0.19	0.85
Asian	0.21	0.76	58	-0.10	0.53	5	1.22	0.29
Chinese	-0.04	0.92	12	0.16	1.15	2	0.24	0.85
Other	0.16	0.64	29	0.02	0.98	7	0.35	0.74
Free School Meals								
Yes	-0.11	0.76	391	-0.14	0.82	61	0.25	0.80
No	0.02	0.72	623	0.05	0.75	103	0.39	0.70
First Language								
English	-0.08	0.75	606	-0.11	0.78	90	0.38	0.70
Other	0.04	0.71	408	0.09	0.76	74	0.54	0.59
Special Educational Needs								
None	-0.01	0.70	815	0.00	0.77	139	0.14	0.89
School Action	-0.29	0.88	82	-0.44	0.92	9	0.49	0.64
School Action Plus	-0.05	0.80	101	-0.03	0.77	14	0.09	0.93
Statemented	0.32	0.81	16	0.54	0.18	2	0.89	0.54