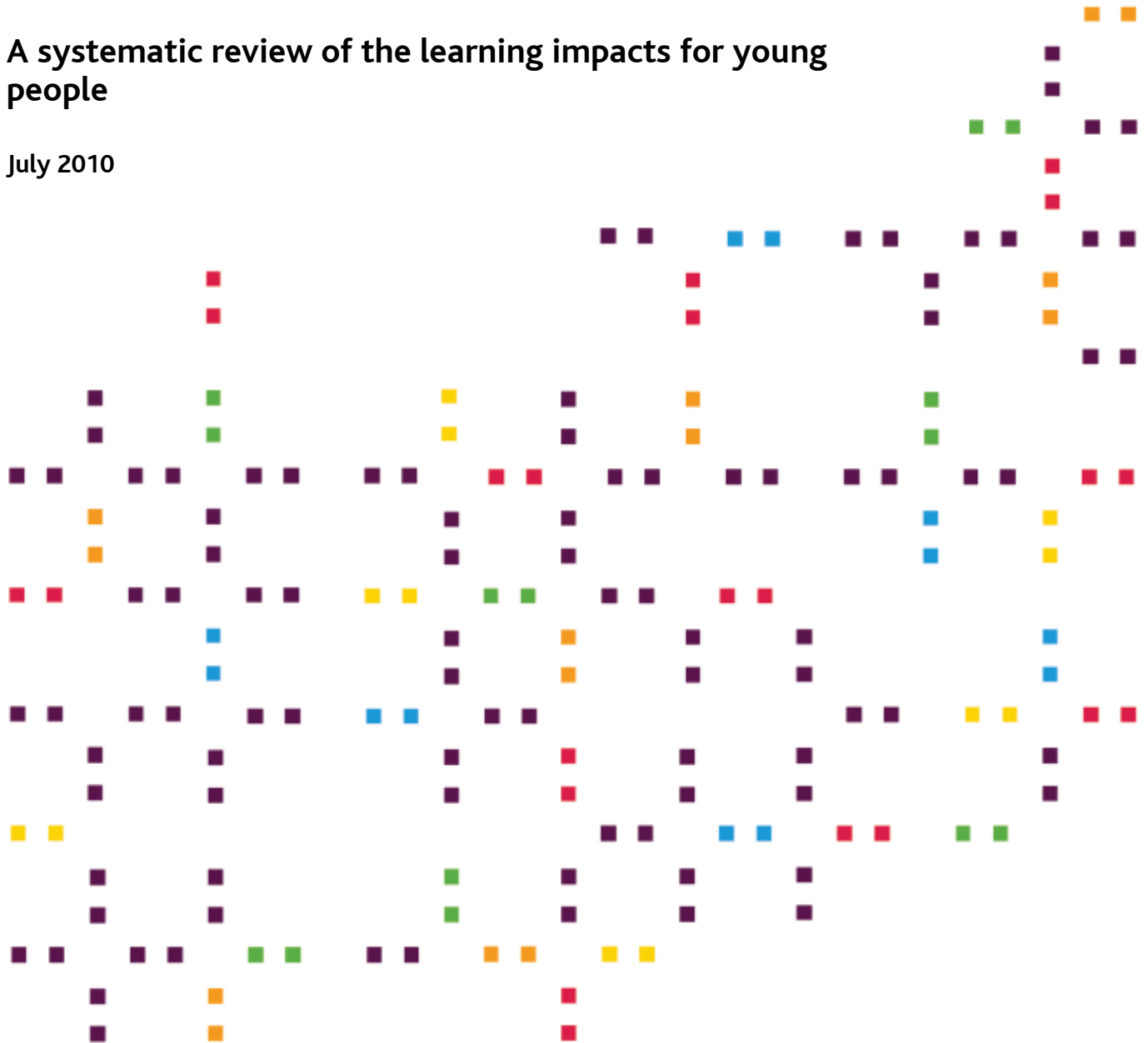




Understanding the impact of engagement in culture and sport

A systematic review of the learning impacts for young
people

July 2010



The EPPI-Centre (Institute of Education, University of London) – The lead contractors and lead on the Impacts work in this report. The work was authored by:

Dr Mark Newman
Dr Karen Bird
Jan Tripney
Naira Kalra
Irene Kwan
Mukdarut Bangpan
Carol Vigurs



This report can be downloaded from the DCMS website:
http://www.culture.gov.uk/what_we_do/research_and_statistics/5698.aspx

Acknowledgements

We would like to acknowledge the support of Chloe Austerberry, Jeff Brunton, Sergio Graziosi and James Thomas for technical support and help with information retrieval.

Contents

| Section | Page |
|--|-------------|
| Background | 6 |
| In-depth reviews of learning impacts for young people... | |
| ...participating in the arts | 14 |
| ...participating in sport | 33 |
| ...attending school libraries | 45 |
| ...attending museums, galleries and heritage sites | 53 |
| Appendices | 62 |
| Arts A1: References | 63 |
| Arts A2: Quality, effect sizes | 68 |
| Arts A3: Summaries | 79 |
| Sports S1: References | 104 |
| Sports S2: Quality, effect sizes | 107 |
| Sports S3: Summaries | 112 |
| Libraries L1: References | 123 |
| Libraries L2: Summaries | 125 |
| Museums M1: References | 137 |
| Museums M2: Summaries | 140 |
| General G1: Systematic review process | 157 |
| General G2: Inclusion and Exclusion Criteria | 158 |
| General G3: Anatomy of a forest plot | 160 |
| General G4: Study quality assessment framework | 161 |
| General G5: Synthesis interpretation framework | 163 |

List of abbreviations

| | |
|--------|--|
| ACE | Arts Council England |
| BME | Black and minority ethnic |
| CASE | Culture and Sport Evidence |
| CaSE-F | Culture and Sport Engagement Framework |
| CAT | Canadian Achievement Test |
| CI | Confidence Interval |
| DCMS | Department for Culture, Media and Sport |
| EH | English Heritage |
| GCSE | General Certificate of Secondary Education |
| KS | Key Stage |
| MSMS | Maryland Scientific Methods Scale |
| MLA | Museums, Libraries and Archives Council |
| NPD | National Pupil Database |
| NDPBs | Non-departmental public bodies |
| SATs | Standard Assessment Tests |
| SE | Sport England |
| SES | Socioeconomic status |
| WoE | Weight of Evidence |
| GPA | Grade point average |
| PE | Physical education |

Background

The Culture and Sport Evidence (CASE) programme was set up by the Department for Culture, Media and Sport (DCMS) in 2008, in collaboration with the sector-leading non-departmental public bodies (NDPBs): Arts Council England (ACE), English Heritage (EH), Museums, Libraries and Archives Council (MLA) and Sport England (SE). The programme aims to generate strategic evidence that will be used to inform the deployment of public funds to maximise engagement in sport and culture, and the value citizens in England receive from that engagement.

As part of the CASE programme, the EPPI-Centre (Institute of Education, University of London) and Matrix Knowledge Group were commissioned to undertake a research project to investigate the drivers, impact and value of engagement in culture and sport. The project used systematic review, analytical and statistical modelling techniques to begin to understand why people engage, or do not engage, in cultural and sporting activities, the benefits they obtain from that engagement, and the potential value to them, and to society as a whole. An overview of the approach and methods of this project can be found in '*Understanding the drivers, impact and value of engagement in culture and sport: Technical Report for the Systematic Review*' published alongside this report on the CASE website.

Aims and rationale for the current systematic review

Empirical research on cultural and sporting engagement is extensive and wide ranging. Through a process of discussion and negotiation, the CASE Board¹ agreed to focus the systematic reviews on a sub-section of this literature: young people's cultural and sporting engagement.

A set of systematic reviews were undertaken to examine the impact of young people's engagement (in each sector: sport, arts, MLA, heritage) on their learning. This systematic review aims to examine the impact of young people's participation in the arts.

A word from CASE

The systematic review described here is a first for our sectors. There have of course been systematic reviews in sports, and in the arts and mental health sectors among others before but this is the first cross-sector review to take the same issue – learning impacts for young people – and review the research available to each area. Why did we choose 'learning impacts for young people'? We chose this area first out of necessity – we needed to make a start on determining what the evidence was for the wider impacts on engagement, and learning impacts for young people represents a priority area. Policies such as the proposed 'school Olympics', and the way libraries, museums and heritage work with education highlight the close connection our sectors have with learning.

The reviews also demonstrate the relative strength of quantitative impact research in each sector: while arts and sports participation domains provided stronger evidence in this area, libraries, museums and heritage research was weaker. While it is clear that these institutions play a core role in providing

¹ The CASE board is the management board responsible for delivering this project. The board members are comprised of the research leads from each CASE organisation.

unique learning opportunities for young people, we need to be better at understanding how and when those impacts happen so that young people of the future have even more rewarding experiences. Better research is key to better understanding what works and how limited resources can be put to best effect.

Finally, the reviews here also signal what kinds of impact research designs are more likely to have more influence with public policy development. We focused on quantitative studies here: principally this focus arises out of a practical need to generate quantitative values for including in policy appraisal (see the accompanying report '*Understanding the drivers, impact and value of engagement in culture and sport: an over-arching summary of the research*' for more on this). There is no question that well-executed qualitative research has a crucial role in developing our understanding of why any impacts detected occur.

We hope that the reviews here and the database that was built as a consequence serve as both an inspiration and a resource for the sectors looking to generate evidence to inform policy of the future.

Methods used in the review

General approach of the project

There is little agreement across the culture and sport sectors about what constitutes 'impact', how or whether impact can be measured, how best to do so and/or whether impacts demonstrated in one context can be generalised to another (Galloway, 2009). Whatever the different positions taken, the justification for any public policy on the topic is that engagement in 'culture or sport' is good for 'something' or 'somebody'. This is a causal claim, i.e., a claim that engagement with cultural or sporting activities produces an effect of some kind, whether it is immediate or long term, direct or indirect. The general approach adopted by this project followed this logic. The in-depth reviews examined research that measured impacts quantitatively and attempted to establish cause and effect relationships between cultural/sporting engagement and an outcome. Decisions about which study designs were appropriate for measuring impact, and methods for assessing quality, were made with the CASE Board. The principle focus on quantitative studies arose out of the desire to generate quantitative parameters that might ultimately be useful for generating value estimates (see 'The Value of Engagement in Culture and Sport: Summary Report' that accompanies this report on the CASE website for more about the value analysis in this project).

User involvement

The Advisory Group, composed of CASE Board members and other relevant stakeholders, played a central role in establishing the conceptual scope of the review outlined above. CASE Board members also decided the focus of the in-depth review questions.

Type of review

This project used systematic review methods to examine the cultural/sporting literature. This approach differs from traditional literature reviews or narrative accounts by using an explicit method to identify, describe and appraise the research studies. A full report of the methods used in the systematic review process can be found in Tripney et al. (2010). The systematic review used standard procedures and processes developed by the EPPI-Centre, as summarised in General Appendix G1 (p. 157).

The review was carried out in three stages:²

Stage one: creating a repository (database) of studies.

The first stage of the review consisted of identifying all research evidence about engagement in culture and sport. It produced a searchable database that is publicly available.

Stage two: mapping exercise

The second stage of the review involved describing (or mapping) a sub-section of the wider literature included in the database created at stage one. The subsection was defined as high quality quantitative studies of impact. High quality studies were defined as those that had a control group and where, in the absence of random allocation to groups, both pre-test and post-test measurements of the outcome of interest were taken. As there were few high quality studies in the MLA and/or heritage sectors, for these two sectors all quantitative impact studies were mapped. Relevant studies were identified and methodological and contextual information was collected. This information provided a basis for informed discussion and decision-making between the research team and the CASE Board about the focus of stage three.

Stage three: in-depth review

The third stage involved an in-depth analysis of a sub-section of the studies mapped in the second stage focusing on impact of learning outcomes for children and young people. Detailed data extraction and quality assessment of the included studies was undertaken and the findings synthesised in order to provide answers to the in-depth review question.

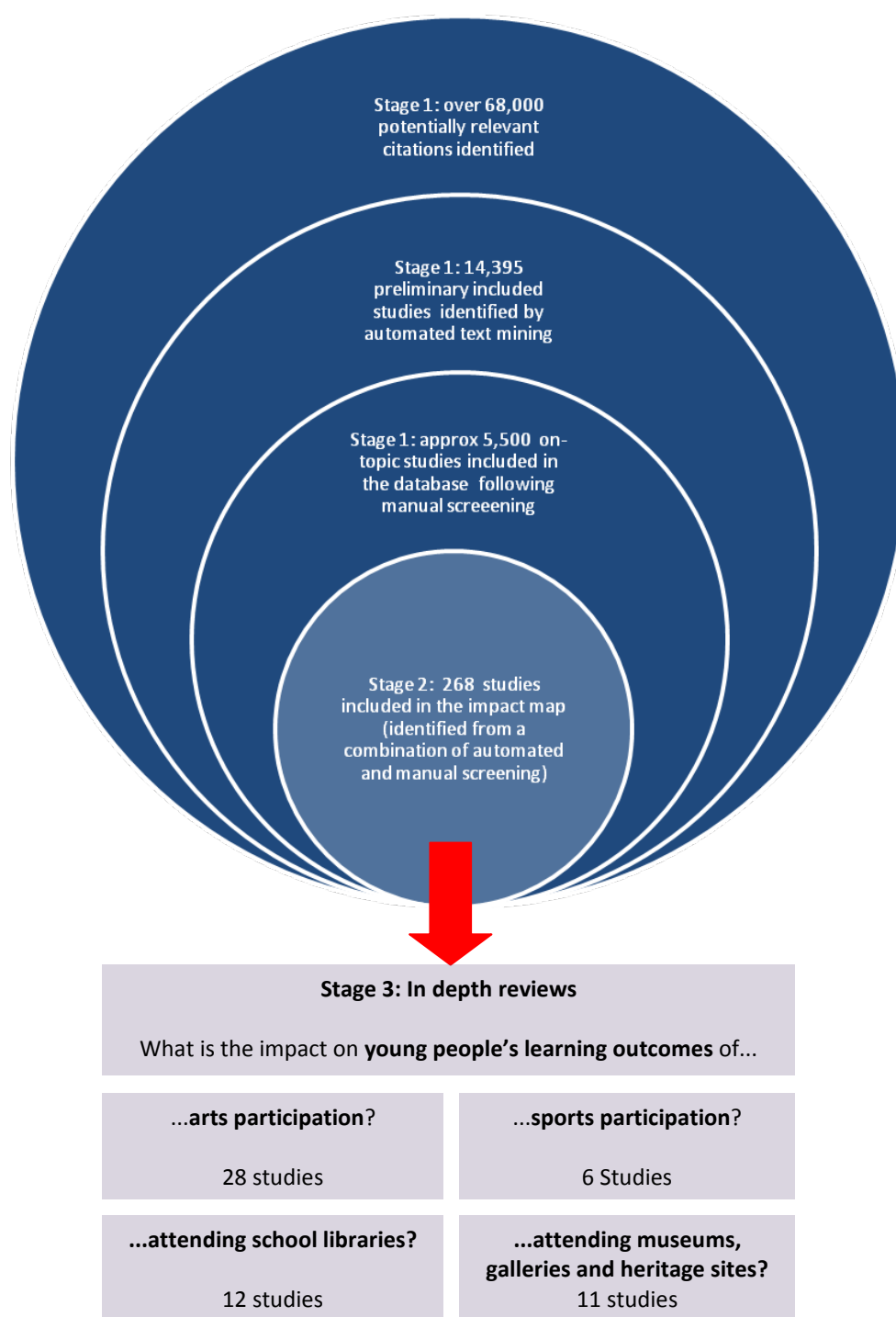
Identifying and selecting studies

A broad and comprehensive search strategy was used to identify research studies on engagement, impact and value in sport, the arts, heritage, museums, galleries, libraries and archives. This generated a very large pool of potentially relevant studies (n=68,434), so text mining software was used to support the process of selection. An overview of the flow of studies through the review is shown in Figure B1.

From mapping of the literature, it became clear that 4 different syntheses would be possible – two based on participation (i.e. Physically engaging in activities) in sports and the arts, two on attendance (i.e. visiting and viewing material) in libraries and in museums, galleries and heritage sites. Within the libraries review, more detailed analysis of the studies revealed that only studies of school libraries formed a coherent group so the CASE board decided to focus that review further still.

² These did not necessarily run sequentially.

Figure B1: Summary of the flow of studies through the different stages of the review process



Inclusion criteria were developed and applied at each stage of the review process (see General Appendix G2, p.158). The in-depth reviews of research on learning outcomes for young people here aimed to include only studies of higher quality by study design. The Technical Report available alongside this report provides a fuller explanation of the process of assessing the quality of studies.

Synthesis for the arts and sports participation reviews

Meta-analytic techniques were used to synthesise the findings from included studies. Each study was subject to an in-depth data extraction, where information about the context, methods and findings were recorded and a quality assessment was carried out (see Arts Appendix A2, p. 68, and Sports Appendix S2, p. 107, for each review). The main aim of the synthesis was to combine results (where appropriate) across studies, therefore we used data that could be translated into a common metric of effect sizes (Hedges' *g*). Most studies reported multiple learning outcomes; where possible, effect sizes were calculated for all outcomes (see Appendices A2 and S2 for details of those calculated for each study in each review respectively).

The in-depth reviews include multiple meta-analyses, each of which addresses a particular sub-question. Studies were grouped according to different types of learning outcome and/or age of participants. For each sub-question, the effect sizes from each relevant study were entered into the meta-analysis to generate an overall pooled estimate of effect. Only one outcome from each study was included in any single meta-analysis (to avoid the same participants' outcomes being counted twice). The individual and pooled effect sizes are displayed on a 'forest plot', which is a diagrammatic representation of the meta-analysis (see General Appendix G3, p.160 for an explanation of how to read a forest plot and the accompanying Technical Report for further explanation).

A minimum of two studies is necessary for a synthesis. Where there was only one study investigating a particular sub-question, brief details about that study have been reported.

Synthesis for the school libraries and museums, galleries and heritage attendance reviews

Due to the different qualities in the literature for these studies slightly different synthesis strategies were used. These are detailed for each review below.

Libraries synthesis

For the purpose of the libraries in-depth review, school library provision was accepted as a proxy for school library attendance as the majority of studies (unless otherwise indicated) investigated the impact of provision, and attendance was assumed. The studies selected for inclusion in the review meet the criterion specified for inclusion in the in-depth review. Only one reasonably high quality quantitative³ study of the impact of school libraries on young people's learning was identified. Therefore all studies that measured the impact of school libraries on young people's learning using some kind of quantitative measure of outcome were included in the in-depth review.

Information and data from each study about context, methods and results was recorded in a process of in-depth data extraction. The results data used were the main quantitative outcome measure of the impact of the intervention (either attendance compared to non-attendance or library provision compared to more developed library provision) on the learning outcomes of the participants. The results used in Table L1 (p. 48) in the libraries in-depth review are based on either the quantitative coefficients (or proportional differences) that measured the relationship between the independent variable (library provision) and the dependent variable (academic attainment) and/or the authors' narrative account of the results. The focus of the synthesis was on making comparisons across and between studies to identify

³ i.e. with a control group and a pre- and post-intervention outcome measure.

what the overall pattern of results tells us about the answer to the question rather than the specific results of any individual study.

The included studies all used approaches in which the statistical association between a measure of academic attainment with quality of provision was analysed. The results of these studies are presented in the form of a coefficient that expresses the direction of the association. Either the slope is positive, indicating that better quality library provision is associated with higher academic attainment, or negative, in which case better quality library provision is associated with lower academic attainment. Synthesis of these studies used a technique known as vote counting. The number of studies with positive or negative coefficients is simply added up and the balance, either positive or negative, is used to draw a conclusion. Unlike the primary study analysis, statistical significance is not considered when deciding whether the findings of a study are positive or negative, as the aim is to identify what the overall body of studies indicates. Statistical significance reflects a number of different issues in study design, including sample size and the variables included in any model. Therefore it cannot be reliably interpreted across and between studies. Furthermore, statistical significance does not provide any indication of practical importance or otherwise.

Museums, galleries and heritage synthesis

The museums, galleries and heritage review identified only one reasonably high quality quantitative study (i.e. Maryland Scientific Methods Scale 3 or above) of the impact of attendance on young peoples' learning. Therefore all studies that measured the impact of museum, gallery and/or heritage site attendance on young people's learning using some kind of quantitative measure of outcome were included.

The information and data from each study about the context, methods and results was recorded in a process of in-depth data extraction. This data was used to carry out the synthesis of the studies.

The studies were organised into two groups, based on differences in the intervention, and synthesised differently:

- 1) studies that examined museums, galleries and/or heritage attendance supplemented by additional learning support (materials or teaching). With the exception of one study, this group of studies had data appropriate for meta-analysis⁴.
- 2) studies that museums, galleries and/or heritage attendance only. This set of studies reported self perceptions of learning outcomes and had data suitable for narrative synthesis.

Meta-analysis

A subset of the studies used research designs and outcome measures that allowed for the findings to be translated into a common metric of effect sizes and so meta-analysed. The main aim of the meta-analysis was to combine results (where appropriate) across studies, therefore we have used data that could be translated into a common metric of effect sizes (Hedges' g). Most studies reported multiple learning outcomes; where possible, effect sizes were calculated for all outcomes reported in these studies.

⁴ Meta-analysis is a statistical technique for combining the results of individual studies to produce a pooled estimate of the direction (positive or negative) and size of effect for a particular intervention. Further information about meta-analysis can be found in the Technical Report published alongside this report on the CASE website: www.culture.gov.uk/case

There are multiple meta-analyses, each of which addresses a particular sub-question. Studies have been grouped according to different types of learning outcome and/or age of participants. For each sub-question, the effect sizes from each relevant study were entered into the meta-analysis to generate an overall pooled estimate of effect. Only one outcome from each study was included in any single meta-analysis (to avoid the same participants' outcomes being counted twice). The individual and pooled effect sizes are displayed on a forest plot which is a diagrammatic representation of the meta-analysis (see General Appendix G3, p. 160, for an explanation of how to read a forest plot).

Narrative synthesis

A second group of studies were not suitable for statistical meta-analysis because they did not use research designs in which outcomes for young people who had attended were compared with those who had not. These studies measured the self perceived impacts of attending cultural institutions. One of these studies also addressed the same question as those studies that could be meta-analysed and thus was also included in the analysis of the question addressed by those studies. The other studies all addressed the same question and measures from all studies were listed, and then cross tabulated to identify where the same or similar outcomes were reported between studies. This was used to create a narrative summary, based on all the outcome measures reported in the different studies.

Drawing conclusions

An adapted version of the Maryland Scientific Methods Scale interpretation framework was used by the review team to interpret the results of the synthesis (Farrington et al., 2002). Details of the framework can be found in the General Appendix G4, p. 161. To illustrate the potential impact of engagement, the pooled effect sizes were transformed into hypothetical changes in relevant test scores (Coe, 2002).

Definitions

The following definitions are used in the reviews:

Young people: refers to anyone under the age of 19 years.

Learning outcomes: refers to the measurement of (but is not necessarily limited to) any of the following:

- academic achievement/skills (as measured by, for example, GCSE examinations or class test scores)
- transferable skills (for example, interpersonal/communication skills, social competency skills)
- cognitive abilities
- truancy rates/behaviour problems
- personal development
- cultural knowledge
- attitude to learning
- capacity or capability to learn/develop

- curiosity
- motivation for learning
- creativity.

Structure of this report

The report is split into two engagement types: reviews focusing on the learning impacts of **participation**, which is limited to the arts and sports sectors, and reviews focusing on the learning impacts of **attendance**, which covered libraries and museums, galleries and heritage. These areas represent the priority areas for the CASE board members. Each report is self-contained and can be read independently of the others, though the same general approach was the same. Individual sector reports may be available from the different CASE members' websites.

Following the in depth review reports are the appendices of supporting information. Each review has its own appendix with up to 3 sections. There is also a general appendix which covers additional information about the systematic review approach. For full details of the methods employed here to conduct the reviews, see the technical report that accompanies this report on the CASE website (www.culture.gov.uk/case).

Learning impacts for young people participating in the arts: an in-depth review

Jan Tripney, Mark Newman, Karen Bird, Carol Vigurs, Naira Kalra, Irene Kwan & Mukdurat Bangpan

Research and policy background

Recent government policies have emphasised the value of maximising the public's engagement in culture and sport and demonstrated a commitment to increasing access to these sectors. This is evident in both national and regional strategies (e.g., *Every Child Matters; A Passion for Excellence*) and across government departments (e.g., *Sea Change; London 2012 Olympic Legacy Action Plan; National School Sport Strategy*). Funded bodies are similarly placing great emphasis on increasing and sustaining participation in culture and sport (e.g., English Heritage's *A Lasting Legacy*; the Arts Council's *Great Art for Everyone*). The current political climate, strengthened by recommendations from the McMaster Review (2008) and recent events (e.g., Liverpool Capital of Culture 2008; preparations for the 2012 Olympics), provides a genuine opportunity for maximising public engagement with culture and sport.

The evidence base for engagement in culture and sport has been reviewed by numerous authors, including projects commissioned by the Department for Culture, Media and Sport (e.g., Daly, 2005; Evans and Shaw, 2004; Freshminds, 2007) and the Scottish Executive (e.g., Galloway et al., 2006; Ruiz, 2004). In the literature, many claims have been made about the impact of the culture and sports sectors on a diverse range of outcomes. These claims, however, are based on a variety of types of knowledge/evidence, and reviews of such knowledge/evidence, having been prepared for different purposes for different audiences, are not consistent in their approach to gathering, selecting, analysing or quantifying the evidence they present. To date, there has not been a comprehensive, systematic, consistent cross-sectoral analysis of the impacts of culture and sport.

In England, Arts Council England (ACE) is the national development agency for the arts, distributing public money from the government and the National Lottery. Between 2008 and 2011, investments will be in excess of £1.6 billion. ACE is accountable to the government and the wider public for the effective use of this public money. As such, it is challenged with understanding the values and benefits that such investments accrue to individuals, communities and society, and with using that understanding to make decisions about the best use of its resources. As part of a commitment to strengthening the existing evidence base on the impact of the arts, Arts Council England and its equivalent in Scotland, the Scottish Arts Council, have funded a number of literature reviews in this area (Arts Council England, 2004; Galloway, 2008; Reeves, 2002), including those looking at the issues of social exclusion and health (Jermyn, 2001, 2004; Staricoff, 2004). The empirical evidence for the impact of the arts has also been reviewed by Coalter (2001), Guetzkow (2002), McCarthy et al. (2001) and Newman et al. (2001), and in reviews examining the impact of arts participation in areas such as regeneration (Shaw, 1999), criminal justice (Hughes, 2005), education (Deasy, 2002; Mason et al., 2006; Standley, 2008;

Winner and Cooper, 2000), social inclusion (Belfiore, 2002) and health (Angus, 2002; Daykin et al., 2008; Gold et al., 2004, 2009; Health Development Agency, 1999; White and Angus, 2003).

Reviews in the field of education have typically focused on children and young people. The contribution of art education to cultural learning in learners aged 5-16 has recently been the focus of a systematic review (Mason et al., 2006), as has the effects of music instruction on children's reading skills (Standley, 2008). There are two substantial surveys of research on the various learning effects of arts participation on children and young people, both conducted by researchers in the United States. To date, the most systematic and comprehensive synthesis in the field of arts education is the research conducted by the Reviewing Education and the Arts Project (REAP), published as a whole issue of the *Journal of Aesthetic Education* (Winner and Cooper, 2000). A total of 188 relevant studies carried out between 1950 and 1999 were identified and a set of 10 meta-analyses conducted. Around the same time, the Washington-based Arts Education Partnership produced a compendium of arts education research in which a qualitative commentary was presented on 62 studies (both qualitative and quantitative research) (Deasy, 2002). These reviews are now quite dated and more recent reviews are limited either by a focus on a specific sub-set of the arts and/or by type of study design.

Both in the UK and internationally, a great deal has also been written about the conceptual and methodological challenges in researching the impact of the arts. There currently remains much disagreement about the meaning of the term 'impact' and appropriate ways of measuring and evaluating it, with many writers questioning the applicability and practicality of the experimental approach for dealing with the complexity of arts impact (Belfiore and Bennett, 2007; Galloway, 2009; Jermyn 2004; Matarasso, 1997). It is argued by some that current methods for assessing the impact of the arts are largely based on a fragmented and incomplete understanding of the cognitive, psychological and socio-cultural dynamics that govern the aesthetic experience (Belfiore and Bennett, 2007).

Recent suggestions for alternative ways of understanding this topic include taking a historical-critical approach (Belfiore and Bennett, 2006) and the use of theory-based evaluation (Galloway, 2009).

Authors, funders, and other users of the review

The project was funded by the Culture and Sport Evidence (CASE) programme⁵. The CASE programme comprises the Department for Culture, Media and Sport (DCMS), Arts Council England (ACE), English Heritage (EH), Museums, Libraries and Archives Council (MLA) and Sport England (SE). The CASE Board is made up of the research directors and/or managers for each organisation, who acted as an Advisory Group.

The project was undertaken jointly by the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London and Matrix Knowledge Group. The EPPI-Centre was responsible for the systematic review work.

Review question

The in-depth review reported here addressed one specific sub-question:

What is the impact of young people's participation in structured arts activities on their learning outcomes?

⁵ <http://www.culture.gov.uk/case/case.html>

What were the findings of the studies?

Descriptive overview of the studies

Twenty-eight studies included for in-depth analysis were data extracted and their quality assessed. Of these, four studies were rated Low or Low/Medium overall weight of evidence and so were not eligible for inclusion in the synthesis (see Section 2.3 for specification of medium/high quality studies).

Of the remaining 24 studies, 15 studies were carried out in the USA, four in Canada, two in Asia (Korea, Hong Kong), one in Germany and two in the UK. **The age range of participants, across all studies, was 3 to 16 years.** In 21 studies, the participants were mixed sex and in one study there were only male participants (in the remaining two studies this information was not stated). Six studies were rated Medium/High overall weight of evidence and eighteen studies were rated Medium. Descriptive summaries for each of the individual studies, including a brief summary of the characteristics of the arts-based intervention, are provided in Arts Appendix A3, p. 79. The results (in the form of the effect sizes calculated for each outcome) can be found in the Arts Appendix A2, p. 68. A full reference list for this part of the report is available at Arts Appendix A1, p. 63.

The 24 studies included in the synthesis investigated the impact of participating in arts activities on young people's learning outcomes.⁶ Of these, three studies measured the impact of arts participation on academic attainment in different subjects (as measured by standardised examinations or tests). Five studies measured the impact on the development of children's early literacy skills (for example, phonological awareness). Eight studies measured changes to children's cognitive skills, and eight studies examined the effect on transferable skills, such as communication skills, social skills and creativity.⁷ A single study investigated whether participation in arts activities improved students' responses to bullying situations and a further single study investigated the impact of an arts activity on pupils' self-protective skills.

Does participation in structured arts activities improve academic attainment?

Overall description of studies

Three studies (two conducted in the UK and one in Canada) investigated the impact of participating in arts activities on academic attainment. Two studies were rated as Medium/High quality (Costa-Giomi, 2004; Kendall et al., 2008) and one was rated as Medium (Fleming et al., 2004). One study included primary and secondary school aged pupils (Kendall et al., 2008) and the other two focused only on primary school children. The families of the children selected to participate in the study by Costa-Giomi (2004) were described as having lower incomes than the typical piano student; a relatively high proportion were single-parent families. The intervention in the study by Kendall et al. (2008) focuses on young people in the most disadvantaged communities in England (where it can be assumed that participants have a different socio-economic profile from young people generally). In the study conducted by Fleming et al. (2004), the primary schools were in the East End of London (Borough of Tower Hamlets) and a large number of the pupils were learning English as a second language.

⁶ Some interventions incorporated attendance at arts-based events.

⁷ Other outcomes may have been reported, but they may not have been relevant to the review (e.g., post-test only) or it may have not been possible to calculate an effect size.

Fleming et al. (2004) used assessments developed by the Performance Indicators in Primary School (PIPS) project based at Durham. These assessments were based on the National Curriculum and widely used by schools across England. The Canadian study by Costa-Giomi (2004) used the Canadian Achievement Test 2 (CAT2). The UK study by Kendall et al. (2008) used national assessment results (SATs and/or GCSEs) for young people reaching the end of Key Stages 2, 3 and 4.

The study by Fleming et al. (2004) examined the impact of the National Theatre's Transformation drama project, which involved attendance/participation in drama activities (writing and performing plays). Kendall and colleagues at the National Foundation for Educational Research (2008) conducted an evaluation of Creative Partnerships, the national creative learning programme in England, designed to develop the skills of children and young people across England through partnerships between schools and creative professionals.⁸ The study by Costa-Giomi (2004) investigated the effects of piano instruction on children's development.

All three studies reported the impact of participating in arts activities on attainment in mathematics and in literacy/English. The study by Kendall et al. (2008) also measured the impact of Creative Partnerships on attainment in science, and in terms of 'best 8' GCSE scores and overall point scores at GCSE and Key Stages 2 and 3. Further information about this study is presented in Box A1.

Initial analysis indicated that the primary and secondary school aged studies were insufficiently similar to combine and thus the synthesis of studies was carried out separately for each group.

Box A1: Evaluation of Creative Partnerships (Kendall et al. (2008))

Creative Partnerships is the national creative learning programme in England, designed to develop the skills of young people, thereby raising their aspirations and achievements. Targeted at the most disadvantaged communities in England, the Creative Partnerships programme brings children and their teachers together with creative organisations and individuals to work in partnerships on creative projects.

The programme supports innovative, long-term partnerships between schools and creative professionals, including artists, performers, architects, multimedia developers and scientists. These partnerships therefore lead to young people's participation in many different types of arts participation activities in different schools.

The study used a matched comparison group design and the National Pupil Database. End of Key Stage 2 and 3 test scores and GCSEs for young people who attended Creative Partnership activities were compared with a matched group of students who did not attend a Creative Partnerships school.⁹

Primary school aged students

All three studies measured outcomes for primary school children. Costa-Giomi (2004) found that students who had three years of weekly individual piano instruction had better language ($g=0.48$, 95%CI 0.11 to 0.85) and mathematics skills ($g=0.30$, 95%CI -0.05 to 0.7) than the comparison group of students who did not get this instruction.¹⁰ Fleming et al. (2004) found that students who participated in the

⁸ <http://www.creative-partnerships.com/>

⁹ Two of authors very generously clarified some issues in order that effect sizes could be calculated.

¹⁰ The author's conclusion that piano instruction did not affect academic achievement in mathematics and language skills appears to have been based on the lack of statistical significance, although this is not entirely clear in the reporting. The use of different statistical techniques can produce different p

National Theatre's three-year drama project performed better at numeracy ($g=0.83$, 95%C.I 0.44 to 1.22) and reading ($g=0.40$, 95%C.I 0.02 to 0.78) than a comparable group of students who did not participate. Kendall et al. (2008) found that, compared to similar young people nationally, primary school students who participated in Creative Partnership activities achieved, on average, the same outcomes in Key Stage 2 tests as those who did not participate.

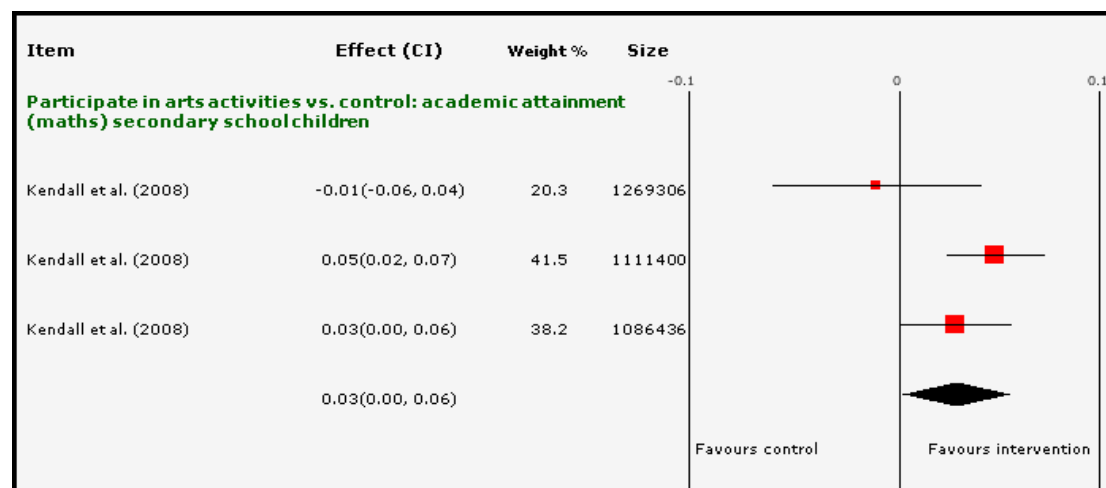
Each of the three studies has results that indicate that primary school aged students who participated in structured arts activities had higher academic attainment outcomes than those that did not. However, statistical analysis suggests that the results of the individual studies are too dissimilar to combine using meta-analysis. On balance, the results would suggest that although arts participation does lead to better outcomes in this age group, there is something about either the type of activity, participants, outcomes or study designs used that requires further investigation. For this reason, we interpret the evidence of impact for primary school aged students as promising, rather than providing evidence that arts participation 'works' for this age group.

Secondary school age students

A single study (Kendall et al., 2008) reported academic outcomes for three different cohorts of secondary school pupils (thus is equivalent to three separate studies). The first cohort of students were those taking SATs at the end of Key Stage 3, the second cohort were students at the end of Key Stage 4 (taking GCSEs in the academic year 2003/4) and the third cohort took GCSEs in the academic year 2005/6.

The impact (for each of the three cohorts) of participating in arts activities on attainment in mathematics is presented in Figure A1, the impact on attainment in English in Figure A2 and the impact on attainment in science in Figure A3.

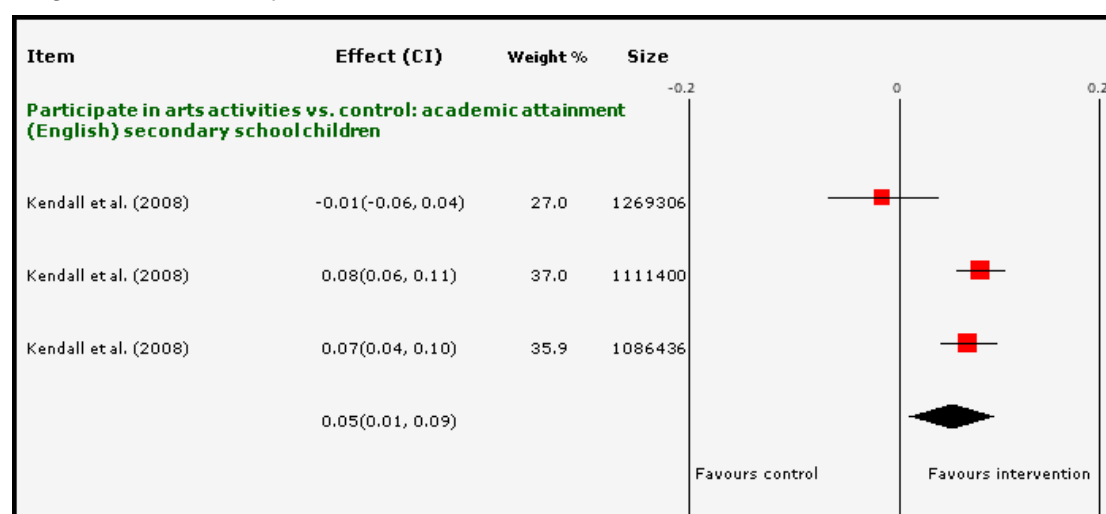
Figure A1: Forest plot illustrating the impact of participating in arts on attainment in mathematics at secondary school level



Heterogeneity statistic $Q = 4.38$ $df = 2$ $p = 0.112$ $I^2 = 54.3\%$, $Q^* = 2.34$; Meta analysis method random effects model

values or confidence intervals, which are used by researchers as the basis for making a claim of statistical significance. The synthesis techniques used in this review do not rely on the statistical significance of any individual study; rather, they are focused on the overall pattern or summary of results.

Figure A2: Forest plot illustrating the impact of participating in arts on attainment in English at secondary school level



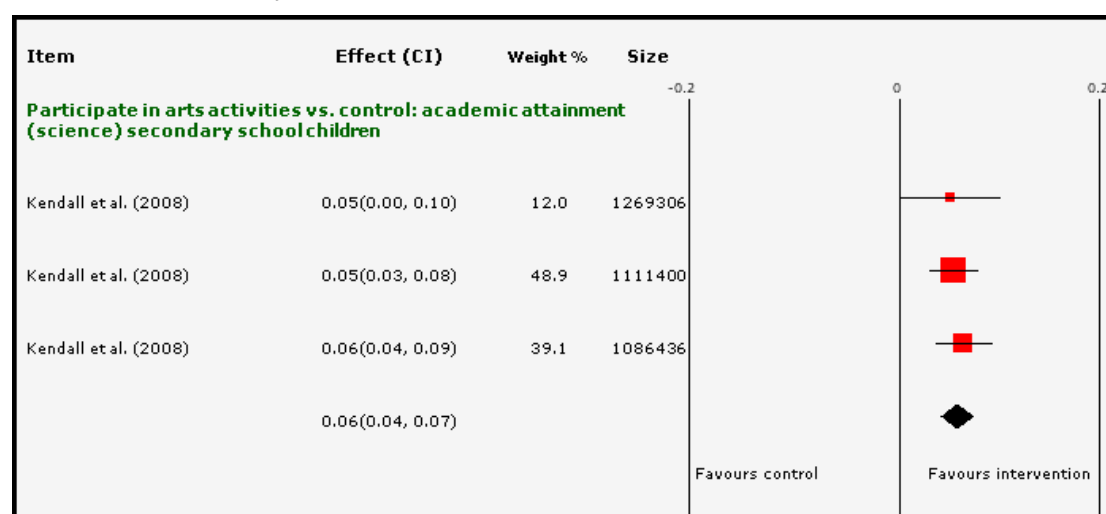
Heterogeneity statistic $Q = 10.9$ $df = 2$ $p = 0.004$ $I^2 = 81.6\%$, $Q^* = 3.12$; Meta analysis method random effects model

For both mathematics (Figure A1) and English (Figure A2) the statistical analysis of heterogeneity indicates considerable difference between these outcomes. However, the difference in the point estimates of effect between each cohort is in total less than $g=0.1$, as are the width of the 95% confidence intervals for each outcome. Furthermore, the total sample sizes are extremely large (this is nearly all driven by the size of the control group, which was comparable pupils in all schools in England taken from the National Pupil Database). With sample sizes this large, even very small differences in the effect sizes between the individual studies will be statistically significant (i.e., suggests statistical heterogeneity). Therefore, we do not think this indicates 'real' systematic differences in the intervention, population characteristics or outcome measures in the different cohorts, but rather does reflect 'real' differences in impact between the different cohorts. Thus, we argue that the pooled estimate of effect can be accepted as a valid estimate of the impact of this intervention on mathematics and English attainment.

The effect sizes for mathematics and English in the initial Key Stage 4/GCSE cohort (2003/4) were negative (i.e., those who did not participate did better than those who did). It is not clear why this should be the case, but it may be important to investigate further whether there were any systematic changes to the Creative Partnerships programme between the earlier and later cohorts that would need to be replicated for the same results to be obtained.

For science, the pooled estimate of effect ($g=0.06$) indicates that the groups that participated in arts activities made greater improvements in academic attainment than those that did not participate in such activities.

Figure A3: Forest plot illustrating the impact of participating in arts on attainment in science at secondary school level



Heterogeneity statistic $Q = 0.364$ $df = 2$ $p = 0.834$ $I^2 = 0\%$, $Q^* = 0.364$; Meta analysis method: random effects model

The overall pooled estimates of effect for all three meta-analyses (as shown in Figures A1, A2 and A3) suggest that the Creative Partnerships programme did lead to increases in academic attainment for those who participated in the scheme. The study was rated Medium/High quality and thus we can be reasonably confident that the differences in progress between the intervention and control groups are valid indicators of the impact of the Creative Partnerships programme. However, although the three cohorts¹¹ are different students, the same students are compared on each of the three subjects (mathematics, English and science). The better results found for those students taking part in Creative Partnerships across all three subjects may suggest that participation in structured arts projects leads to a general improvement in attainment rather than something subject specific.

It is difficult, however, to decipher the exact mechanism that produced this improvement in attainment, as the description of the Creative Partnerships scheme is not very detailed and it would appear that there was considerable scope for different activities under the umbrella of the Creative Partnerships programme, which may or may not be an important element of the programme's success.

Does participation in structured arts activities improve early literacy skills?

Five Medium quality studies conducted in the USA measured the impact of arts participation on the development of children's early literacy skills (Gromko, 2005; McMahon et al., 2003; Piro and Ortiz, 2009; Register, 2004; Standley and Hughes, 1997). The studies used different instruments to measure different aspects of early literacy. The studies by McMahon et al. (2003) and by Piro and Ortiz (2009) focused on children of primary school age (grades/years 1 and 2), while the remaining three studies considered impacts on younger children.

One study evaluated the effectiveness of an arts-based educational programme, 'Basic Reading Through Dance' (McMahon et al., 2003). Piro and Ortiz (2009) evaluated keyboard instruction. The remaining three studies evaluated multi-component music interventions that incorporated activities such as singing, playing instruments and movement (Gromko, 2005; Register, 2004; Standley and Hughes, 1997). In all five studies, the main goal of the intervention was to enhance early

¹¹ Key Stage 3; Key Stage 4/GCSE: 2003/4; Key Stage 4/GCSE: 2005/6.

literacy skills (four studies focused on reading and one on writing). Each individual lesson therefore involved the fusion of both arts-based activities (music or dance) and literacy activities. An example of how these two different types of activities were combined in a single initiative is provided in Box A2.

Box A2: Evaluation of Basic Reading Through Dance (McMahon et al., 2003)

The 'Basic Reading Through Dance' programme was a curriculum-based reading intervention developed by Whirlwind, a not-for-profit organisation in the US. This arts-based educational programme had one general goal: to improve students' early-reading skills. Each session in the programme required students to use their bodies to physically represent the alphabet symbols for various sounds in the English language, as well as combinations of sounds. To physically represent each sound, students had to (a) visualise the appropriate symbols (letters) for spoken sounds and then (b) recreate these images physically using their bodies. Three trained artists implemented the programme, which was delivered in schools over 10 weeks (twice a week in 40-minute sessions).

The study used an unmatched comparison group design and participants in this study were first-grade students from Chicago public schools. The study examined the effects of the programme on reading skills, as assessed by different sub-tests of Read America's PhonoGraphix Test.

Table A1: Outcomes and effect sizes, multi-component arts interventions for early literacy

| | Letter naming <i>g</i> (95%C.I) | Phonemic tests <i>g</i> (95%C.I) | Other <i>g</i> (95%C.I) |
|----------------------------|--|---|---|
| Gromko (2005) | 0.1 (-0.3 to 0.5) | 0.7 (0.3 to 1.1) | Nonsense word fluency -0.3 (-0.7 to 0.05) |
| McMahon et al. (2003) | Vowels 0.36 (0.2 to 0.52) Consonants 0.6 (0.43 to 0.75) | 0.6 (0.4 to 0.8) | |
| Register (2004) | 0.2 (-0.4 to 0.8) | | Reading -0.2 (-0.8 to 0.4) Fluency 0.3 (-0.3 to 0.9) |
| Piro and Ortiz (2009) | | | Vocabulary 1.28 (0.85 to 1.7) Verbal sequencing 1.5 (1.1 to 2.0) |
| Standley and Hughes (1997) | | | Writing and language skills 0.9 (0.2 to 1.6) |

These five studies appear to evaluate similar interventions in the same age group. However, the instruments used vary, as does the vocabulary used by the study authors to describe the different outcomes they measure (see Table A1 for details). Whilst it is clear that these outcomes are all measures of a global phenomenon of

'early literacy', it is not clear whether the different instruments/outcomes are measuring the same or different sub-phenomena. Whilst it is possible to combine subsets of the studies that are not statistically heterogeneous, we do not find any of such combinations plausible or convincing. Therefore, we have not reported a pooled effect size for this group of studies. Narrative numerical synthesis using vote counting indicates that ten outcomes from five studies found that those students who participated in the interventions had better 'early literacy' performance scores than those who did not participate, and three outcomes were inconclusive. This compares to two outcomes from two studies where those children who participated in the arts intervention had worse 'early literacy' performance scores than those who did not.

Does participation in structured arts activities improve cognitive abilities?

Eight studies investigated the impact of participating in arts activities on cognitive abilities, five in the USA and three in Canada. Two studies were rated Medium/High quality (Costa-Giomi, 2004; Schellenberg, 2004) and six were rated Medium (Bilhartz et al., 1999; Gromko and Poorman, 1998; Orsmond and Miller, 1999; Rauscher et al., 1997; Rauscher and Zupan, 2000; von Rossberg-Gempton et al., 1999). Studies used a number of different standardised testing instruments. Two studies included primary school aged children (Costa-Giomi, 2004; von Rossberg-Gempton et al., 1999) and the remaining six studies focused on younger children.

One study evaluated the effectiveness of a creative dance programme (von Rossberg-Gempton et al., 1999). In the remaining seven studies, the type of arts activity was music. Two studies (Bilhartz et al., 1999; Gromko and Poorman, 1998) evaluated multi-component music interventions that incorporated activities such as singing, playing instruments and movement. Four studies investigated the impact of keyboard/piano instruction (Costa-Giomi, 2004; Rauscher et al., 1997; Rauscher and Zupan, 2000; Schellenberg, 2004), of which two conducted additional separate evaluations of the impact of singing on children's cognitive skills (Rauscher et al., 1997; Schellenberg, 2004). Orsmond and Miller (1999) evaluated the impact of a Suzuki music programme for classical music instruction. The study by Schellenberg (2004) also evaluated the effectiveness of drama lessons.¹²

Box A3: Evaluation of keyboard, voice and drama lessons (Schellenberg, 2004)

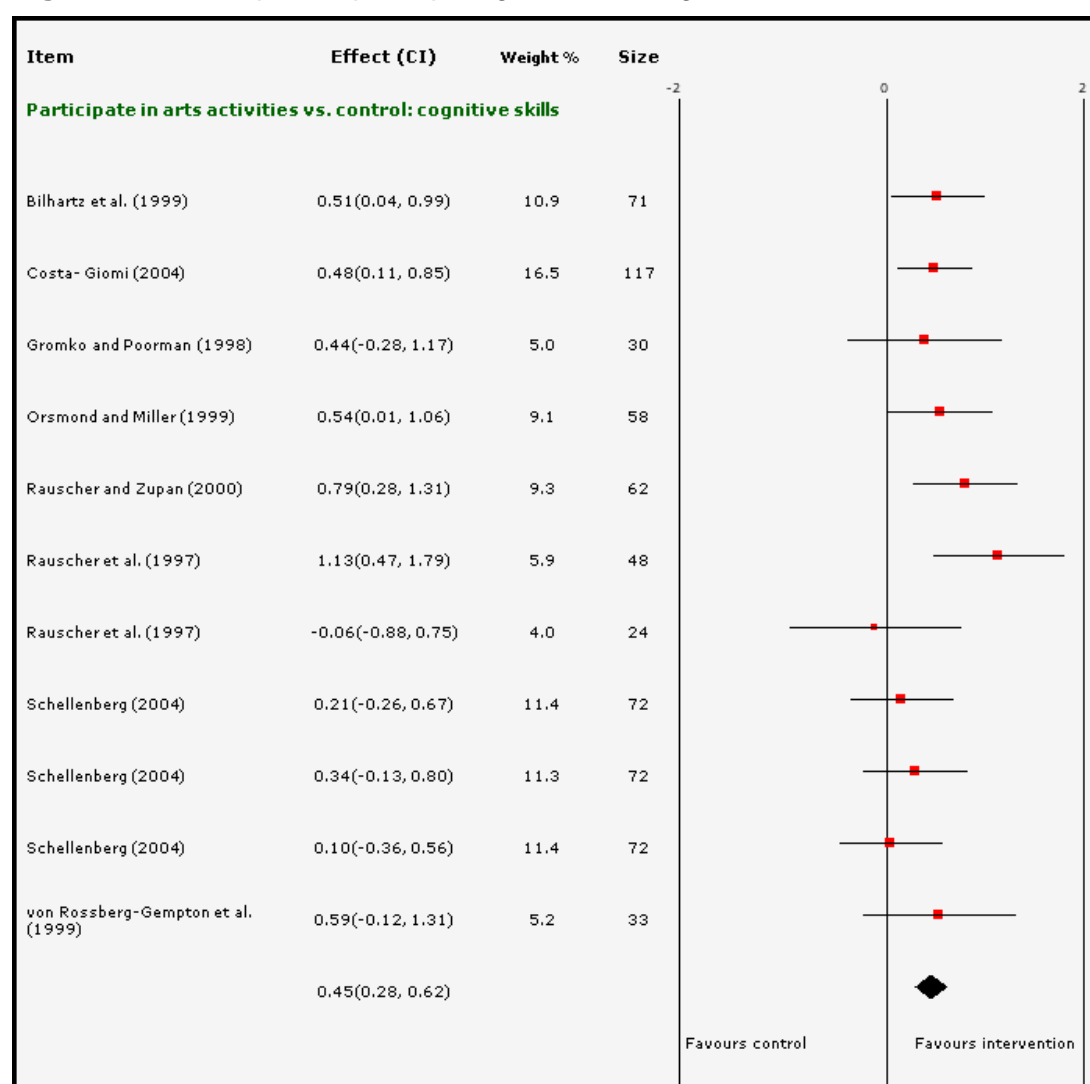
This study evaluated the effects of weekly arts lessons on children's IQ. The lessons were taught for 36 weeks at the Royal Conservatory of Music (Toronto, Canada). Lessons involved practice and rehearsal, memorisation, learning new scripts or pieces, expressing emotion, and so on.

The author reports that children were randomly allocated to one of four groups (36 children in each group): two groups received music lessons (keyboard or voice), one group received drama lessons and one group received no lessons. Participants were recruited at the age of six years. The author reported that six-year-olds were selected because children of this age are considered to be sufficiently mature for formal lessons, and because absolute (perfect) pitch is evident among individuals who begin taking music lessons before the age of seven.

¹² In the study, students participating in this activity served primarily as a control group.

Statistical analysis suggested that these studies are similar enough to combine using meta-analysis. Figure A4 shows a forest plot of the effect sizes for each individual study and the pooled estimate of effect. The pooled estimate suggests that young people who participated in arts activities made greater improvements in cognitive performance than the group that did not participate in such activities. The results also suggest that the size of effect may not be the same for all types of structured arts participation. However, this is complicated by the variation in the intensity and direction of the participation investigated in each study.

Figure A4: The impact of participating in arts on cognitive abilities



Heterogeneity statistic $Q = 11$ $df = 10$ $p = 0.34$ $I^2 = 10.3\%$, $Q^* = 10.1$; Meta analysis method: random effects model)

Does participation in structured arts activities improve transferable skills?

A group of studies (eight in total) investigated the impact of arts participation on what might collectively be called transferable skills, including self-constructs, communication skills, social skills and creativity. A number of the studies measured more than one of these outcomes. It is not clear, however, whether these outcomes are separate phenomena or different manifestations of the same underlying phenomenon. The results for each are reported separately in the sub-sections below. We feel, however, that the results should best be interpreted as a whole. Taken together, we argue that the results suggest that participation in structured arts

activities does lead to improved transferable skills compared to non-participation in such activities.

Does participation in structured arts activities improve self-constructs?

Four studies (three carried out in the USA and one in Canada) investigated the impact of participating in arts activities on students' self-constructs: i.e., views about oneself that predict motivation and performance to varying degrees. One study was rated Medium/High overall weight of evidence (Freeman, 2001) and three were rated Medium (Catterall, 2007; Catterall and Peppler, 2007; von Rossberg-Gempton et al., 1999). Three studies included only primary school children and one study (Catterall, 2007) investigated outcomes for students attending middle schools.

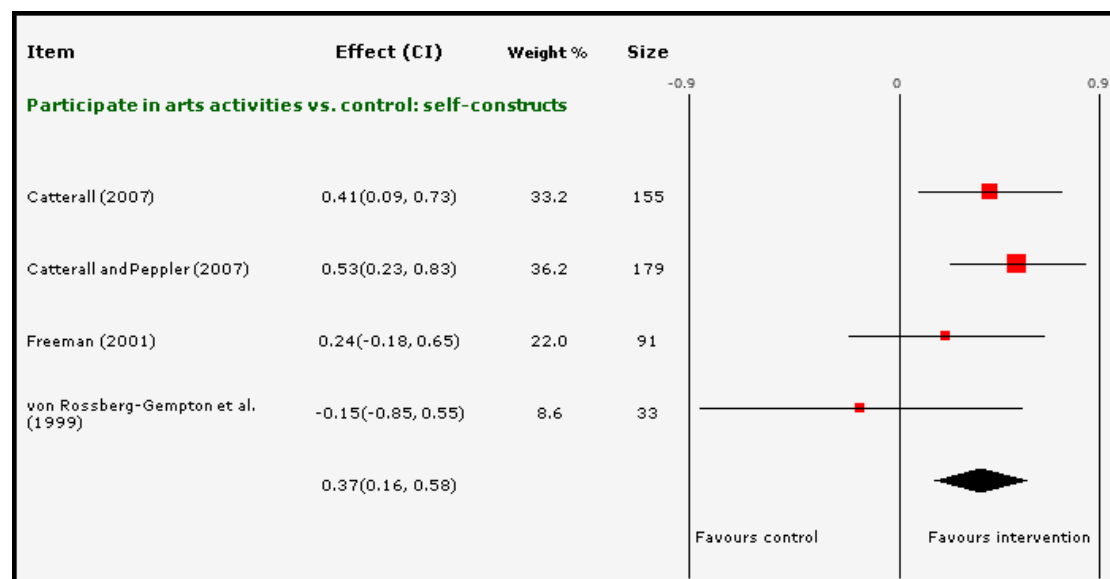
The study by Catterall (2007) examined the impact of a school-based extra-curricular multi-component drama initiative called the 'School Project', which used a range of activities, including attending a professional play, as building blocks for students to write and perform plays (for further details see Box A4). The dissertation study by Freeman (2001) investigated the impact of multi-component creative drama activities within a school context. Catterall and Peppler (2007) investigated the effects of a multi-component visual arts initiative on inner city children in two major US cities. At one site, the activities were drawing, painting and sculpting, and at the other site the programme involved creating visual art and writing poetry. The final study in this group (undertaken as the main author's dissertation) evaluated the effectiveness of a creative dance programme (von Rossberg-Gempton et al., 1999).

Box A4: Evaluation of the 'School Project' (Catterall, 2007)

This North American drama programme used theatre, movement, writing, voice, drawing, and visual arts exercises as building blocks for students to write and perform original plays. Students participated in 90-minute extra-curricular workshops once a week, for 24 weeks. During the course of the programme, all students were taken to see a professional play. The programme took place at three middle-school sites in Los Angeles. Schools were in areas impacted by crime, drug-trafficking and economic hardship. More than 80 percent of participants were from low-income families, and two of the three school sites were among the lowest academic performing schools in the city school system.

The programme aimed to develop students' sense of self, strength in standing up for one's own views and ideas, respect for the views of others, and more positive social interactions. A range of transferable skills were measured, including self-efficacy.

The statistical analysis of heterogeneity suggests that the studies are sufficiently similar to combine. The pooled estimate of effect ($g=0.37$) indicates that students who participated in arts activities had, on average, higher self-construct scores than those who did not (see Figure A5). However, the confidence intervals of two studies cross zero, including that of the one Medium/High study (Freeman, 2001). One of the four studies also reported a negative effect (von Rossberg-Gempton et al., 1999). It is not immediately clear why the results in this study should be negative. One possibility relates to the enrolment of participants in the control group in a specific exercise programme, whereas in the other studies the control groups received 'treatment as usual' (so, in effect, the same comparison was not being made in the von Rossberg-Gempton study as in the other studies).

Figure A5: The impact of participating in arts on self-constructs

Heterogeneity statistic $Q = 3.6$ $df = 3$ $p = 0.302$ $I^2 = 17.7\%$, $Q^* = 3.16$; Meta analysis method: random effects model

Does participation in structured arts activities improve communication skills?

Three studies (one conducted in the USA and two in Asia) investigated the impact of participating in arts activities on communication skills. One study was rated Medium/High overall weight of evidence (Kim et al., 2008) and two were rated Medium (Bigelow, 1997; Hui and Lau, 2006). In one study, the participants were pre-school aged boys with autism (Kim et al., 2008). Hui and Lau (2006) focused on children of primary school age and Bigelow investigated impact on high school students. Conducted in Korea, the study by Kim et al. (2008) investigated the impact of participating in music activities (improvisational music therapy) on children's non-verbal social communication skills. The dissertation study by Bigelow (1997) examined the impact of a school-based performing arts course (involving instrumental and vocal music, drama and dance) on students' communication skills. Hui and Lau (2006) conducted their study in Hong Kong, evaluating the impact of an extra-curricular multi-component creative drama project on children's communicative-expressive ability (see Box A5 for further details).

Box A5: Evaluation of multi-component drama education (Hui and Lau, 2006)

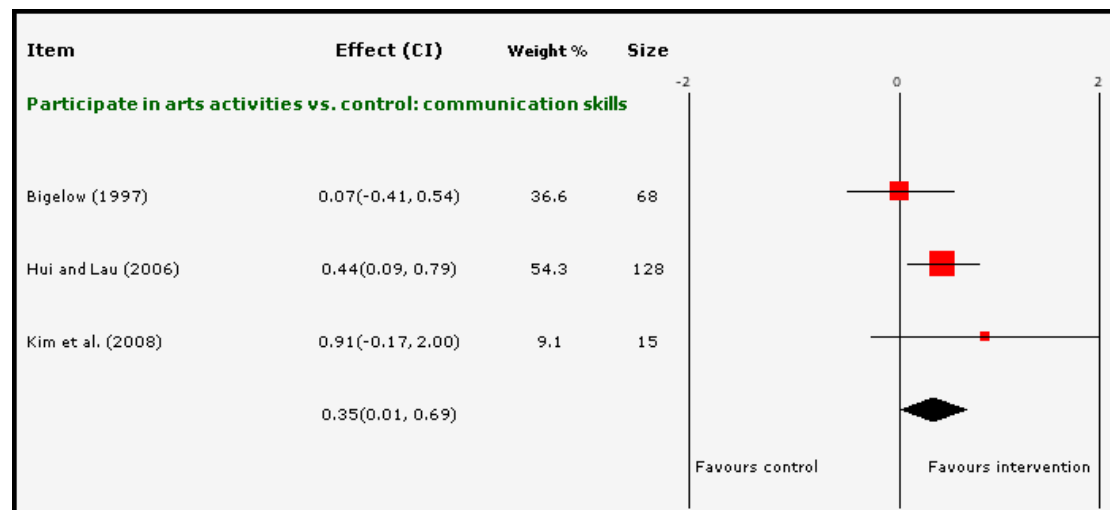
The multi-component drama project evaluated in this study involved puppet making and creative drama (improvisation and story creation). The project was jointly developed by a professional theatre group for children and a non-government organisation for children. The curriculum of the project incorporated knowledge, techniques and attitude in children's development of aesthetics, and creativity, as stated in the Curriculum Guidelines for Primary School Education by the Education Department (Hong Kong). The programme was delivered after school, one day a week for 16 weeks.

An unmatched comparison group design was used. Seventeen elementary schools participated in the drama project. The control group participated in non-aesthetic extra-curricular activities. Changes in students' communicative-expressive ability and creativity (thinking and drawing) were measured.

The statistical assessment of heterogeneity suggests that the studies are similar enough to combine and the pooled estimate of effect ($g=0.35$) indicates that the

group that participated in arts activities made greater improvements in communication skills than the group that did not participate in such activities (see Figure A6). However, caution is required in interpreting these results. The study by Kim et al. (2008) found a comparatively large effect size and was conducted with a very small sample with different characteristics from the samples in the other two studies (in this study the participants were all male and had autism). When this study is removed from the meta-analysis (forest plot not shown), the pooled estimate of effect is smaller ($g=0.29$) and is inconclusive (95% C.I. -0.06 to 0.64).

Figure A6: The impact of participating in arts on communication skills



Heterogeneity statistic $Q = 2.62$ $df = 2$ $p = 0.27$ $I^2 = 23.6\%$, $Q^* = 2.07$; Meta analysis method: random effects model

Does participation in structured arts activities improve social skills?

Four studies (two conducted in the USA, one in Canada and one in Asia) investigated the impact of participating in arts activities on social skills. Three studies were rated Medium/High quality (Freeman, 2001; Kim et al., 2008; Wright et al., 2006) and one Medium (Catterall, 2007). In one study, the focus was on pre-school aged boys with autism (Kim et al., 2008). Two studies investigated the impact of arts participation on primary school children (Freeman, 2001; Wright et al., 2006). Catterall (2007) focused on children of secondary school age.

Kim et al. (2008) investigated the impact of participating in music activities (improvisational music therapy) on children's pro-social behaviours. Catterall (2007) examined the impact of an extra-curricular multi-component drama initiative called the 'School Project', which used a range of activities, including attending a professional play, as building blocks for students to write and perform plays; pro-social changes in behaviour were measured (for example, problem resolution skills). Freeman (2001) investigated the impact of school-based multi-component creative drama activities on pupils' social skills.

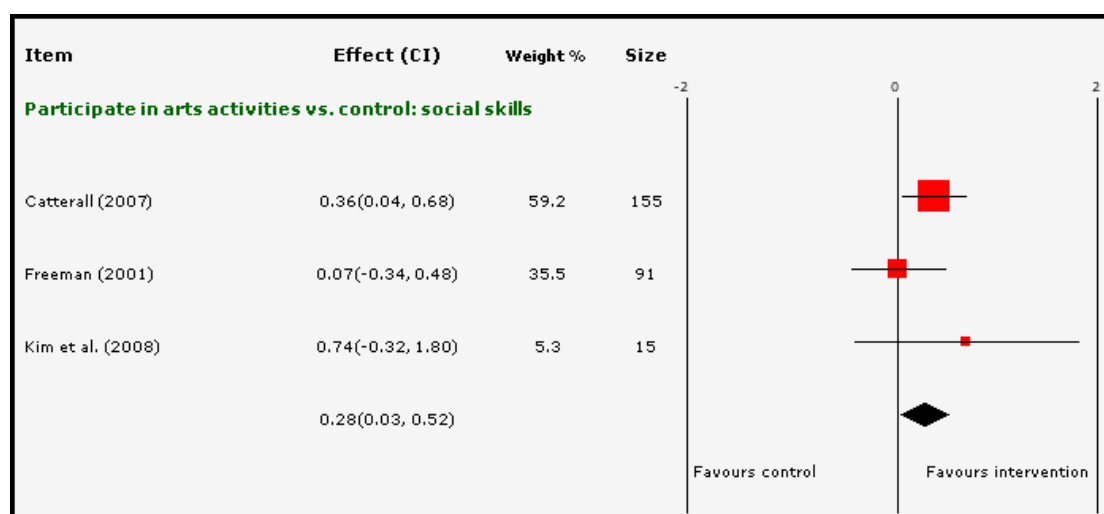
Wright et al. (2006) examined the impact of a structured multi-component arts programme targeted at low-income communities (the National Arts and Youth Demonstration Project) on a range of outcomes for young people, including conduct and pro-social behaviour (see Box A6 for further details). The analysis of this study used hierarchical linear modelling and the outcomes were beta-coefficients. With the data reported, it was not possible to calculate standardised mean difference effect sizes. However, the results suggest that the change in scores for participants and non-participants across all the outcomes measured was for all practical purposes the same, i.e., there was no difference in outcome between the groups.

Box A6: Evaluation of a structured arts programme (Wright et al., 2006)

This study evaluated a structured arts programme for Canadian youth, aged 9 to 15 years, from low-income communities. The aim of the programme was to engage youth in productive and life-enriching activities rather than correct or treat problems. Young people engaged in structured arts instruction (involving a combination of theatre, visual and media arts) as part of an after-school arts programme delivered twice a week over nine months. Sessions were 90 minutes long. The study used a well-matched comparison group design involving 183 young people. A control group was selected using propensity score matching, with data drawn from the National Longitudinal Survey of Children and Youth (NLSCY).

Excluding Wright et al. (2006), the analysis of statistical heterogeneity suggests that the studies are similar enough to combine (see Figure A7). The pooled estimate of effect ($g=0.28$) indicates that the groups that participated in arts activities made greater improvements in social skills than the group that did not participate in such activities. Nonetheless, caution is required in interpreting these results. The study by Kim et al. (2008) found a larger effect size than the other two studies and was conducted with a very small sample with different characteristics from the samples in the other two studies (in this study the participants were all male and had autism). When this study is removed from the meta-analysis (forest plot not shown), the pooled estimate of effect is smaller ($g=0.24$) and less conclusive (95% C.I. -0.03 to 0.52). Furthermore, as noted above, the study by Wright et al. (2006) found no difference in effects between participants and non-participants.

Figure A7: The impact of participating in arts on social skills



Heterogeneity statistic $Q = 1.97$ $df = 2$ $p = 0.373$ $I^2 = 0\%$, $Q^* = 1.97$; Meta analysis method: random effects model

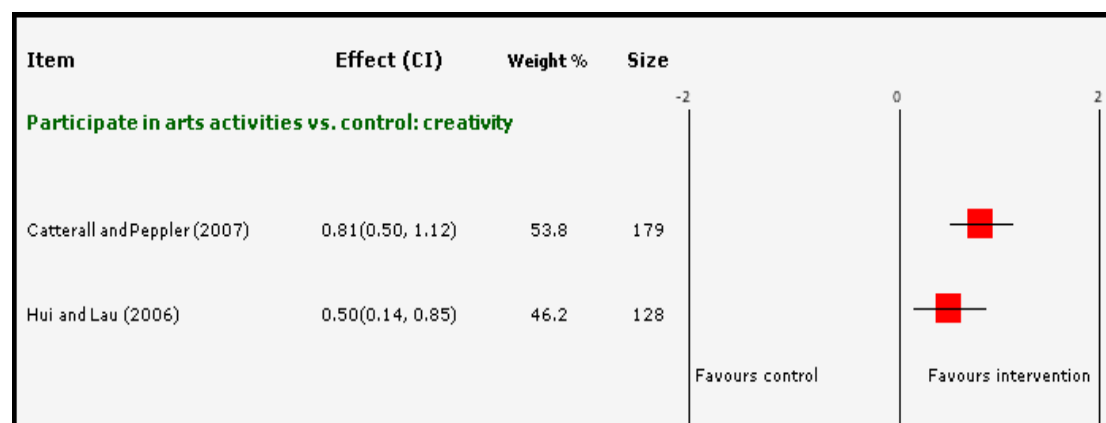
Does participation in structured arts activities improve creativity?

Two studies (one conducted in Asia (Hong Kong) and one in the USA) investigated the impact of participating in arts on young people's creativity. Both studies were rated Medium quality and were conducted with children of primary school age. Hui and Lau (2006) evaluated an extra-curricular multi-component creative drama project involving puppet making, improvisation and story creation. Catterall and Peppler (2007) investigated the effects of multi-component visual arts instruction on inner city children in two major US cities. At one site, the activities were drawing, painting and

sculpting, and at the other site the programme involved creating visual art and writing poetry.

In both studies, the effect size was positive and excluded 'no difference', suggesting that those who took part in the arts activities made greater progress in developing creativity. However, the statistical analysis of heterogeneity suggests that the studies are too dissimilar to combine using meta-analysis (see Figure A8).

Figure A8: The impact of participating in arts on creativity



Heterogeneity statistic $Q = 1.75$ $df = 1$ $p = 0.186$ $I^2 = 42.8\%$, $Q^* = 1$; Meta analysis method: random effects model

Other study results

Two studies not yet discussed answered a further two different sub-questions. Synthesis was not possible, as each sub-question was addressed by only one study.

A single study investigated whether participation in arts activities improved students' responses to bullying situations (Merrell, 2005). It was rated Medium overall weight of evidence. This North American doctoral dissertation examined the effectiveness of a drama-based social skills programme (the 5 W's of Bullying Intervention) with a population of 56 urban high school students in the ninth grade. For two outcomes (attitudes to bullying and direct intervention in bullying situations), the direction of effects was positive but inconclusive; for the third outcome (change in reporting bullying) the effect was negative and inconclusive.

A single study investigated the impact of an arts activity on pupils' self-protective skills (Krahe and Knappert, 2009). It was rated Medium overall weight of evidence. The study was described as a group-randomised trial of a theatre-based intervention to prevent sexual abuse. The intervention (No Child's Play) was targeted at first and second grade primary school children in Germany. It also included three-hour training sessions for teachers to prepare them for their task in guiding the children through the performance and a three-hour information evening for parents designed to provide facts and raise awareness about sexual abuse. One hundred and forty-eight students participated in the study. The group who watched a live performance of a play had better performance in the target variables (distinguishing good/bad touch, secrets, getting help, and rejecting unwanted touch) than the control group who did not watch the play ($g=0.70$).¹³

¹³ The first author very generously provided additional data in order that effect sizes could be calculated.

Discussion and implications

Implications for policy and practice

Research evidence does not on its own provide justification for any particular policy decision. Thus, the implications provided here are simply the review teams' interpretations of what we think the balance of evidence indicates is the answer to a particular question using the interpretation framework.

What works:

When compared to non-participation in structured arts activities:

Participation in structured arts activities improves academic attainment in secondary school aged students. Participation in such activities could increase their academic attainment scores by 1% and 2%, on average, above that of non-participants (all other things being equal).

Participation in structured arts activities improves pre-school and primary school aged children's early literacy skills. This result is based on narrative numerical synthesis and thus we are unable to estimate the size of any positive effect.

Participation in structured arts activities improves young people's cognitive abilities (based on various measures of intelligence). Participation of young people in such activities could increase their cognitive abilities test scores by 16% and 19%, on average, above that of non-participants (all other things being equal).

Participation in structured arts activities improves young people's transferable skills. Participation of young people in such activities could increase their transferable skills test scores by 10% and 17%, on average, above that of non-participants (all other things being equal).

What is promising:

There is insufficient yet promising evidence about the impacts of participation in structured arts activities on primary school aged children's academic attainment. The three studies that investigated this question were insufficiently similar to combine using meta-analysis. However, the preponderance of evidence was positive.

What is unknown:

There is insufficient evidence about the impacts of participation in structured arts activities on young people's responses to bullying situations. However, this does not mean that these impacts do not occur.

There is insufficient evidence about the impacts of structured arts participation and/or attendance on young people's self-protection skills. However, this does not mean that these impacts do not occur.

There are two specific issues about the research evidence included in this review that should be taken into account. Firstly, the issue of 'compared to what' and, secondly, the issue of the 'independence' of the subjects in the studies. In most studies, the comparison group either received 'no treatment' or an alternative non-arts activity; however, this does not necessarily mean that they were not engaging with the arts at all. (In some studies, the comparison group would appear to have received 'treatment as usual', e.g., their standard arts lesson at school.) For the studies where the outcomes measured were academic attainment, the statistical analysis of the data carried out in the synthesis assumes that all the participants are

completely independent of one another. However, the intervention and control participants were allocated at the school level, i.e., all intervention participants were in participating schools and all control group participants were in non-participating schools. This means that the results obtained by individual participants at any particular school are not completely 'independent'. This 'clustering' usually has the effect of making results less conclusive.

There are, of course, many other issues that need to be considered when making judgements about what policy actions, if any, these interpretations suggest. Wider considerations include knowledge of policy and practice in the field, knowledge about economic costs, knowledge about values and beliefs, knowledge about other desirable and undesirable consequences of (in)action, and knowledge about other potential means for achieving the same ends. Any such discussion would also need to consider the generalisability of interventions and study results from the context in which they were generated into the context in which they are used. Discussion of this type is not the role of this review and, as such, the findings presented here form the beginning rather than the end of any investigation and analysis in this area.

Implications for research

One of the aims of this review was to highlight potential research directions suggested by the current evidence base in this area.

Secondary research

The systematic map (conducted in stage two) identified research evidence that had quantitative measures of the impact of engagement in cultural/sporting activities. In all, 92 high quality studies were included in the map. This in-depth review has only examined a small part of that evidence base (studies that focused on the impact of participating in arts activities on young people's learning outcomes). Future research may wish to explore additional areas of the systematic map. In-depth reviews could be conducted, for example, on the impact of arts participation on different population groups (e.g., older people or women only) or for different outcomes (e.g., social outcomes). This in-depth review only synthesised high quality studies. Further in-depth reviews could also incorporate lower quality studies into the synthesis to further explore the findings reported here.

Further relevant and interesting review work could endeavour to compare the impacts of the interventions included in this review with other potentially relevant policy alternatives. For example, a meta-analysis that assessed the effects of personal development for teachers on student academic outcomes found an overall effect size of $d = 0.66$ (Timperley et al., 2007). Whilst effect sizes obtained in different meta-analyses (e.g. arts participation and CPD for teachers) are not directly comparable, analysis of such indirect comparisons could provide further useful information for policy decision-making.

Primary research

The in-depth review identified 24 high quality studies that examined the impact of young people's participation in arts activities on their learning outcomes. This is a relatively small yield, suggesting that further primary research is required.

In order to develop the knowledge base, studies need to use designs that adequately control for bias and have sufficiently large populations to facilitate 'transfer' into policy and practice. There also appears to be little, if any, cost-effectiveness analysis in this area, therefore such analysis should be a feature of any new evaluation research that is commissioned.

Strengths and limitations of this systematic review

This systematic review, undertaken as part of the 'understanding the drivers, impact and value of engagement in culture and sport' project, is, as far as we are aware, the first to have attempted to be comprehensive, systematic and transparent across such a wide body of literature in the field of culture and sport.¹⁴ As such, both the database and the individual in-depth reviews provide an important resource for the culture and sport communities, not only in their content but also in the development of systematic review methods for future investigation of the questions that remain unanswered in the field. However, the review represents only the first step in an ongoing process of building knowledge and understanding about the impact of engagement in the arts. The in-depth review addressed only a very small part of the agenda of interest but, importantly, the systematic and comprehensive approach used means that it will be possible to utilise the database of studies to begin to address some of the other questions of interest in subsequent reviews.

However, there are limitations to the review processes that should be taken into account when considering the results that are presented here. There is a detailed presentation and discussion of the limitations of these processes in the review technical report (Tripney et al., 2010) and only a summary is provided here. The general approach taken by the project was to provide evidence that could be used within a decision-making framework that assessed the relative impacts and values of investments in different kinds of cultural and sporting activities. A different approach may have led to different results. The scope of the review (in terms of what is defined as 'arts') was limited to those arts listed in the 'Taking Part' survey.¹⁵ The process of initial searching, whilst comprehensive, was limited to studies published in the English language, and some studies identified in the searches could not be obtained (mainly US dissertations). The use of text mining technology facilitated the initial broad scope of the review but may have led to relevant studies not being identified in the selection process.

Table A1: Results of the quality assurance exercise

| | Relevant | Missed studies | | | Consequences |
|---------------|---------------------------------------|------------------------|---------------------------|-----------------------|---|
| | Relevant studies (total) ¹ | Missed studies (total) | Missed by search strategy | Missed by text-mining | Missed studies relevant to the in-depth review ² |
| Hallam (2009) | 59 | 44 | 34 | 10 | 3 |

1. Items in this column were identified by screening studies (based on title) in the reference list of the review against the database inclusion criteria

2. Items in this column were identified by screening the missing studies against the in-depth review inclusion criteria. (This could only be completed for those studies for which we were able to obtain full texts; for nine studies, reports were unobtainable.)

Some limited quality assurance of the search and selection processes was undertaken to assess the extent to which studies may have been missed, either in the initial search or by the text mining. We looked at the reference lists of a recent review of arts participation to identify studies that, based on their title, would appear

¹⁴ There are systematic reviews on particular aspects of the field, such as factors influencing sports participation and impacts of the arts.

¹⁵ http://www.culture.gov.uk/what_we_do/research_and_statistics/4828.aspx

to have been relevant to this review and thus should have been in our database. The results of this exercise are shown in Table A1.

The results suggest that our initial search did not identify a number of relevant studies in each sector. A large proportion of the missed studies were North American dissertations, or other grey literature, which are not routinely included in electronic databases and which may have been obtained by the review authors through personal contacts. Some missed studies, however, were published in peer reviewed academic journals and it is not yet clear why these were not detected by our search strategy. Ten studies were not identified by the text mining, all of which were about some aspect of music participation. A likely explanation for this is that the studies, in their titles and/or abstracts, did not use a general descriptive term like 'music' but used a specific term like 'choir'. If the sample of included studies used to create the text mining search string did not include a study that used the specific term in its title or abstract, then this term would not be in the text mining algorithm and thus any studies using only this term would not be identified as relevant to the review.

All relevant items identified during the quality assurance exercise were manually entered into the database. The three available studies that were identified as answering the in-depth review question were included in the synthesis of arts impact studies detailed in this report.

The selection of studies for the in-depth review was limited to only high quality (based on the Maryland Scientific Methods Scale) quantitative studies. It could be argued that this is too restrictive a definition of 'impact'. Whilst defensible, it is acknowledged that this approach may have led to the exclusion of some studies that were 'as good as' those studies that were included. The boundaries between levels two and three of the Maryland Scientific Methods Scale are not necessarily as clear cut as the scale itself might suggest.

The learning impacts for young people participating in sport: an in-depth review

Karen Bird, Mark Newman, Jan Tripney, Carol Vigurs, Naira Kalra, Mukdarut Bangpan & Irene Kwan

Policy and research background

Cross sector

Recent government policies have emphasised the value of maximising the public's engagement in culture and sport and demonstrated a commitment to increasing access to these sectors. This is evident in both national and regional strategies (e.g. *Every Child Matters*; *A Passion for Excellence*) and across government departments (e.g. *Sea Change*; *London 2012 Olympic Legacy Action Plan*; *National School Sport Strategy*). Publicly funded bodies are similarly placing great emphasis on increasing and sustaining participation in culture and sport (e.g. English Heritage's *A Lasting Legacy*; The Arts Council's *Great Art for Everyone*). The current political climate, strengthened by recommendations from the McMaster Review (2008) and recent events (e.g. Liverpool Capital of Culture 2008; preparations for the 2012 Olympics), provide a genuine opportunity for maximising public engagement with culture and sport.

The evidence base for engagement in culture and sport has been reviewed by numerous authors, including projects commissioned by the Department for Culture, Media and Sport (e.g., Daly, 2005; Evans and Shaw, 2004; Freshminds, 2007) and the Scottish Executive (e.g., Galloway et al., 2006; Ruiz, 2004). In the literature, many claims have been made about the impact of the culture and sports sectors on a diverse range of outcomes. There has not been, however, a comprehensive, systematic, consistent and cross-sectoral analysis of the impacts of culture and sport. This series of in-depth reviews begins to address this gap. Systematic methodology is used to review and synthesise research on the impacts of cultural and sporting engagement. This report focuses on one of the in-depth reviews: the impacts of sporting engagement.

Sports policy background

Sports participation has become a policy priority for the government. With the forthcoming Olympic Games, increased funding for sporting organisations and ambitious targets for increasing participation, the government is committed to maximising public engagement in sport. Since 1997, the government have recognised the value of sport for engendering social inclusion and neighbourhood renewal. In 1998, a Policy Action Team was established to identify the contribution that sport (and the arts) could make to tackling social exclusion.

The report from this group, commonly known as PAT 10, highlighted the positive contributions that sport could make to health, crime, education and employment outcomes (see Policy Action Team 10, 1999). Following this, the government published *Game Plan* (DCMS, 2002), a document widely recognised as a landmark in sports policy development. *Game Plan* established the government's aim to

increase and widen the base of participation in sport and physical activity. This strategy was underpinned by recognition that sporting participation can assist neighbourhood renewal and deliver a range of social benefits.

Alongside this, sporting organisations responded to the social inclusion agenda by developing strategies to increase sports participation. Sport England developed a strategic direction for sport through research and impact evaluation, setting out ambitious targets for increased participation in the *Framework for Sport in England* (2004). This has been superseded by the *Sport England 2008-2011 Strategy* (2008b), which includes targets for 1 million people to do more sport by 2012-13. The Youth Sport Trust took on a central role in delivering high quality physical education (PE) and promoting sports participation in schools (Tacon, 2008). A variety of policies were also developed to target participation by particular groups, such as women and Black and Minority Ethnic (BME) groups (UK Sport, 2003; UK Sport 2004)

In 2008, the government published *Playing to Win* (DCMS, 2008) and restated the importance of sports participation. This report, however, marked a change in direction. Compared to an emphasis on social exclusion and the wider benefits of sport, *Playing to Win* recognised the intrinsic value of sport and acknowledged the reasons why most people participate. Particular importance was placed on enabling people to access and benefit from competitive sport.

Sports research background

Over the past ten years, the UK research agenda has mirrored the policy interest in sport and social exclusion. The literature has examined the social and economic impacts of sporting engagement, establishing the value of sport for individuals and society.

Complementing this social inclusion agenda, a stream of research has examined the nature and level of sporting participation for different groups. This has highlighted low levels of sporting participation in particular socio-demographic groups (such as women and BME groups); they are, therefore, less likely to experience the benefits of sporting engagement.

This has prompted further research examining the barriers to sporting participation (e.g., Foster et al., 2005; Long et al., 2009). The sports participation literature is diverse, including different approaches and research methods. The main bodies of research, primarily with a UK focus, are briefly described below. This growing body of primary and secondary literature is indexed in the Value of Sport Monitor,¹⁶ a database of up-to-date research and critical reviews.

A number of reviews have focused on the social and economic impacts of sport. Playing an influential role in policy discussions, Collins et al. (1999) highlighted the contribution that sport could make to the social inclusion agenda. This was followed by a set of reviews by Coalter and colleagues (2000; 2001; 2003; 2004; 2005) that examined the social benefits of sport and its potential role in regeneration. At the same time, Ruiz (2004) established an evidence base for public policy by reviewing national and international research on the social and economic impacts of sport (and culture more generally).

Studies have examined the potential for sport to improve communities and reduce crime. Sport England's series of guides *Shaping places through sport* (2008a) sought to examine the relationship between sport and the wider outcomes associated with strong and vibrant communities. The role of sports participation in crime reduction and prevention is currently not established in the literature. This is partly a reflection of the complexity of the issue.

¹⁶ http://www.sportengland.org/research/the_value_of_sport_monitor.aspx

Sport may produce a set of positive individual outcomes that lead to social outcomes through a series of intermediate processes. These processes are not yet fully understood (Tacon, 2008). There is evidence to suggest that sporting engagement contributes to community cohesion. Individual studies point to the social networking and social capital outcomes of sport. For example, statistical analysis for the DCMS identified a strong correlation between sports participation and social trust (Delaney and Keaney, 2005). This report highlighted a variety of social benefits associated with being a member of a sports club.

The literature highlights a range of direct and indirect economic benefits associated with sports participation. The UK sporting economy has witnessed high levels of growth over the past twenty years, generating sport-related employment and consumer expenditure (Sport Industry Research Centre, 2007). Economic evaluations of sporting events also highlight the economic benefits for local areas in terms of employment, expenditure and regeneration (Ruiz, 2004).

Beyond the social and economic benefits, research highlights the health and educational outcomes of engaging in sport. The health benefits of sport and physical activity are widely accepted (Chief Medical Officer, 2004; Powell and Pratt, 1996). These pertain to both physical and mental health benefits (Fox, 1999; Scully et al., 1998). Whilst most of the research relates primarily to physical activity, there is a growing body of evidence that examines the links between sport and health (see Tacon, 2008; Galloway et al., 2006). Recent research considers the important role played by sport in preventing and tackling childhood obesity, as well as delivering a range of other health and well-being outcomes (Aarnio et al., 2002; Tacon, 2008; Ekland et al., 2005).

Several reviews have examined the relationship between physical activity and educational performance. These identify evidence that physical education and activity can improve educational outcomes (Sallis et al., 1999; Shephard, 1997). However, these reviews do not employ comprehensive search strategies, assess the quality of included studies, or provide estimates of the size of any impacts. There is little research that examines the specific mechanisms through which sport influences educational attainment (Tacon, 2008).

Authors, funders, and other users of the review

The project was funded by the Culture and Sport Evidence (CASE) programme).¹⁷ The CASE programme comprises the Department for Culture, Media and Sport (DCMS), Arts Council England (ACE), English Heritage (EH), Museums, Libraries and Archives Council (MLA) and Sport England (SE). The CASE Board is made up of the research directors and/or managers for each organisation, who acted as an Advisory Group.

The project was undertaken jointly by the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London and Matrix Knowledge Group. The EPPI-Centre was responsible for the systematic review work.

Review question

The in-depth review reported here addressed one specific question:

What is the impact of young people's participation in organised sport on their learning outcomes?

¹⁷ <http://www.culture.gov.uk/>

What were the findings of the studies?

Descriptive overview of the studies

Six studies evaluated 'sports interventions' and measured 'learning' outcomes for young people aged 4-16 (See Sports Appendices S3 (p. 112) for a descriptive summary, and S2 (p. 107) for quality rating for each study). A full reference list for this review is given at Sports Appendix S1 (p.104). It is important to note that the majority of these 'sports interventions' are non-sporting activities that take place within a sporting context. The details of the interventions are reported in the results sections below. Descriptive examples of the types of study included in the review are provided in Boxes S1-S4.

Four studies measured the impact of participation in sport on academic achievement (numeracy or literacy) and four on 'transferable skills' such as self-efficacy. Other learning outcomes included capability to learn and truancy/behavioural problems. The quality of the studies ranged from medium (n= 3) to medium high (n= 2), with one study having both medium and medium high rated outcomes.

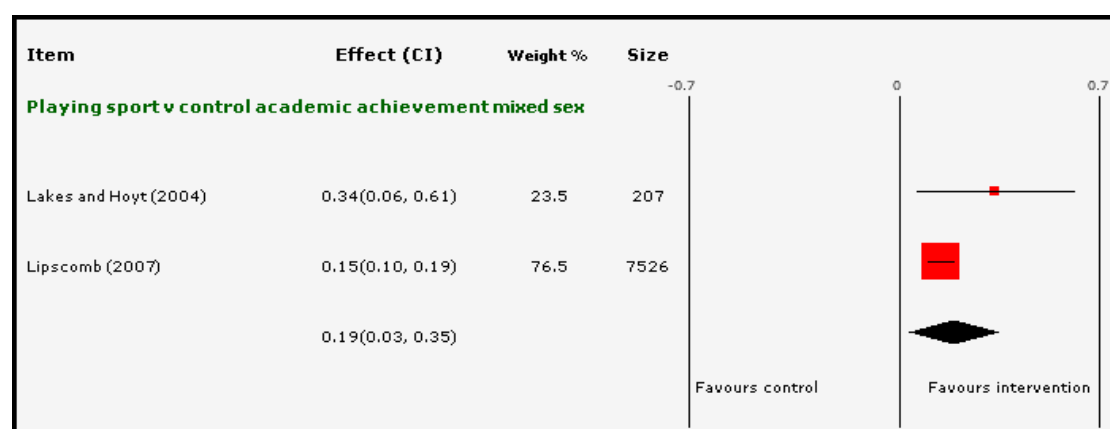
Two of the studies were carried out in the UK and the other four studies were undertaken in the USA. The UK interventions were carried out within a school setting (one in primary school and the other in primary and secondary). Two of the US interventions were delivered within schools (one primary and one secondary) and two were carried out in other settings (these are unknown as the setting was not clearly reported by the authors). The majority of the studies included mixed sex participants, with two studies including only male participants in their sample. The UK studies focused on underachieving young people or those at risk of school exclusion. One of the US studies included a population with limited English proficiency. Two of the remaining US studies included a mainly white population and one did not report details of the participants.

All of the studies in the review made reference to theories about how participation in sport could improve learning outcomes. In the Lipscomb (2007) study this was not translated into a specific hypothesis; rather, several 'plausible explanations' for how participating in sport could improve learning outcomes were put forward. Four of the studies specifically formulated hypotheses that articulated the mechanisms through which participating in sport promoted improved learning outcomes – through promoting self esteem, self regulation, and social competence – i.e., these 'capacities' are promoted through participation in sport and are necessary for successful learning (Anderson 1999, Freitag 2006, Lakes and Hoyt 2004, Powell et al. 2008). In 'Playing for Success', the sports participation (football coaching in professional football clubs) is hypothesised to be a factor that motivates young people to undertake the linked educational programme that is provided as part of the intervention (Sharp et al. 2003a).

Does participation in organised sport increase young people's academic attainment?

Based on two studies (Lakes and Hoyt, 2004; Lipscomb, 2007), this synthesis found that students who played organised sport achieved higher numeracy levels than students who did not play organised sport. The effect sizes from each study, and their combination ($g = 0.19$), are mapped out on the forest plot below (see Figure 3.1). The heterogeneity statistics (representing how comparable the studies are) indicate that the studies are dissimilar. However this may be an artefact of the large difference in the size of the samples in the two studies (207 and 7526 students respectively). Therefore we judge that the pooled estimate of effect represents a valid estimate of the average impact of playing sports on academic achievement as indicated by these two studies.

Figure S1: Forest plot of the impact of participation in organised sports activities on academic attainment (compared to non participation)



Heterogeneity statistic $Q = 1.78$ $df = 1$ $p = 0.182$ $I^2 = 44\%$, $Q^* = 1$, Meta-analysis method: random effects model

Both Lakes and Hoyt (2004) and Lipscomb (2007) were rated as medium/high quality (see General Appendix G4, p.161 for details of how these judgements were reached). The effect size in Lakes and Hoyt (2004) is larger and with a wider confidence interval than Lipscomb (2007). This difference may be a reflection of the different type of sporting activities and/or population groups examined. Alternatively, these effect sizes may represent the minimum and maximum likely estimates of the impact of playing sport on academic achievement.

Both of the studies were carried out in the USA and investigated the impact of participation in sport on numeracy. Lakes and Hoyt (2004) examined the impact of Taekwondo on primary age students and Lipscombe (2007) assessed the effectiveness of a range of sports on secondary school students (see Boxes S1 and S2 for further details).

Box S1: Evaluation of secondary school extra-curricular sports (Lipscomb,2007)

This US study assessed the impact of extra-curricular activities on short term learning outcomes in numeracy and science (which are correlated in the longer term with labour market success). The authors used a sample of 16,305 students from the National Educational Longitudinal study of 1988 (NELS). The empirical analysis used data from the baseline year (a population of 8th graders) to compare with outcomes from the student population in the subsequent two years' samples. The authors controlled for exogenous factors

that may affect participation in extracurricular activities as well as factors such as gender, 'race' and school type.

Box S2: Evaluation of Leadership Education through Athletic Development (LEAD) (Lakes and Hoyt, 2004)

This US study evaluated a school-based martial arts training programme called Leadership Education Through Athletic Development (LEAD). The programme took place within the USA and trained students in Taekwondo. The authors hypothesised that participation in the LEAD programme would prompt an increase in student self-regulation, which is associated with an increased capacity for success and a decrease in self-destructive behaviour. A randomised controlled trial was undertaken to measure the effect of students' participation in Taekwondo on self regulation. To measure this, the authors used a range of outcomes from academic attainment to capacity to learn.

193 primary school students were randomly assigned to the intervention or control group. Data was collected before the students participated in LEAD and then three months later.

Do extra-curricular learning activities linked to organised sport improve academic attainment for underachieving pupils?

'Extra-curricular learning activities linked to organised sport' refers to activities that take place within a sporting setting or following a sporting activity. These interventions are NOT sporting activities but activities that accompany some sport-related activity. Therefore, sport is not the primary mechanism for delivering the learning outcome but rather provides a site or setting for the delivery of extra-curricular learning. Two studies evaluated this kind of intervention (see Boxes S3 and S4 for details of each study).

Box S3: Evaluation of 'Promoting Achievement Through Sports' (Freitag, 2006)

'Promoting Achievement Through Sports' (PATs) was an after-school sports programme designed to provide children with positive experiences and opportunities to learn life skills. Taking place in the USA, the programme combined football training sessions with life skills educational classes. The participants attended PATs twice weekly, with football matches played on Saturdays. Freitag (2006) measured the impact of PATs on the academic achievement (measured by Grade Point Average - GPA) of low achieving Latino boys.

Box S4: Evaluation of 'Playing for Success' (Sharp et al., 2003a)

'Playing for Success' was a national UK sports initiative directed at underachieving primary and secondary school pupils from urban areas. The programme aimed to improve literacy, numeracy and ICT skills by establishing study support centres in local football clubs or other sports venues. These were made available to young people outside school hours. Students that attended the programme were supported by volunteer mentors. The duration of the programme was typically 20 hours spread over 10 weeks. Playing for Success ran for a number of years with different schools/centres.

This study assessed the effectiveness of the Playing For Success in its fourth

year. The evaluation was carried out by the National Foundation for Educational Research. A total of 1,132 pupils, 351 parents and 91 teachers took part in the Spring of 2002. Many of the 58 centres that took part were newly established since the previous year's evaluation. The study measured the effect of Playing For Success on young people's literacy, numeracy and ICT skills. Numeracy and literacy skills were measured by nationally standardised tests that were designed for the study. ICT skills were reported by the participants via a self-completion questionnaire.

The main finding of this synthesis is that students who took part in the extra-curricular activities improved their academic attainment (numeracy) more than a pool of comparable students who did not take part in these extra-curricular activities (see Figure S2). This finding requires the following qualifications:

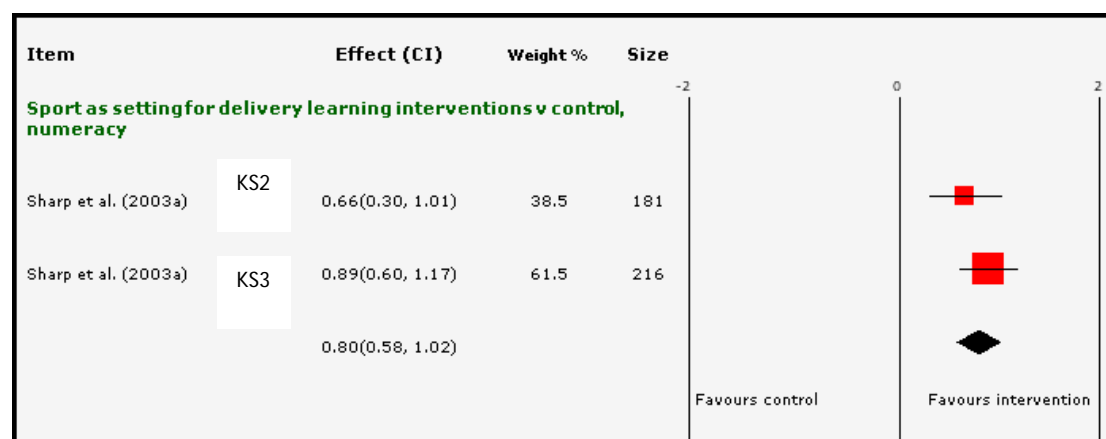
- The finding is based on effect sizes calculated from one study only (Sharp et al., 2003a) and brings together numeracy outcomes as recorded at Key Stage 2 (primary) and 3 (secondary), with a pooled effect size of $g = 0.80$. Combining the outcomes from the two different studies (Sharp et al., 2003a and Freitag, 2006) was not possible because the analysis suggests that there are important differences between these studies. Analysis suggests that the heterogeneity may not be due to differences in interventions, ages or study quality between the studies/outcomes, but rather, that the different measures of academic attainment may be important. Sharp et al. (2003a) measured numeracy outcomes via national standardised tests and Freitag (2006) measured academic attainment using GPA (cross-subject grades).
- Whilst the study (Sharp et al., 2003a) was graded as medium/high, there are limitations with the reporting that mean that we should be cautious when drawing inferences from the observed positive effect. The findings of the national evaluation are reported across a number of different publications. This presents difficulties for the synthesis as different sample sizes and different student cohorts are reported across the research reports. This is problematic because the synthesis requires an accurate reporting of sample size (student numbers in the intervention and control groups) and student cohorts (drawing comparisons between students of the same year/cohort).
- It is not clear how important, if at all, the sporting context is for achieving the learning outcomes reported here. It would appear that 'sport' provided an incentive for the children to undertake the learning activities. It is not clear whether the learning activities would have been undertaken had a different incentive been used.

Do extra-curricular learning activities linked to organised sport improve ICT skills for underachieving pupils?

The Sharp et al. (2003a) study also reports outcomes for ICT skills. These outcomes are reported for two separate student groups: students at Key Stage 2 (primary) and students at Key Stage 3 (secondary). Individually, the outcomes for each Key Stage suggest that students who took part in the extra-curricular activities improved their ICT skills more than a pool of comparable students who did not take part in these extra-curricular activities (shown in Figure S3). However, it is not possible to bring these findings together using meta-analysis (combining outcomes for Key Stage 2 with Key Stage 3) because there is too much heterogeneity. This explains why there is no pooled effect size on Figure S3. The heterogeneity may be a reflection of the different impact of the intervention on different population groups. It is likely that

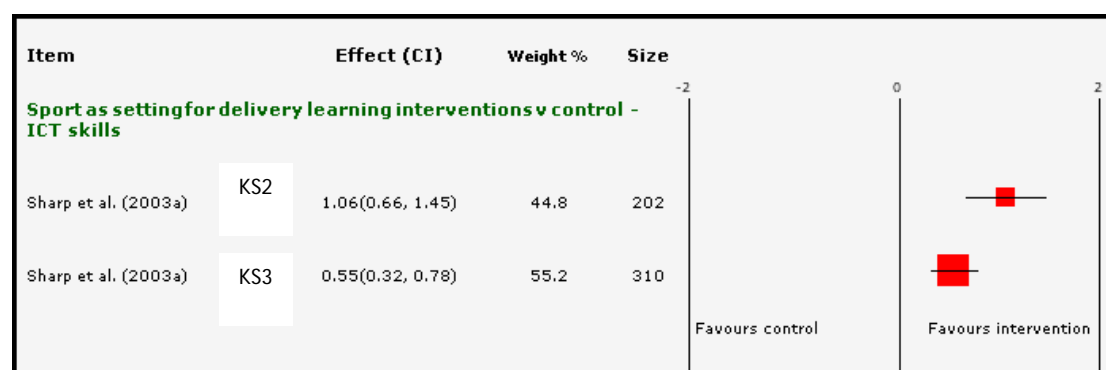
primary and secondary school students have different levels of exposure to ICT. Moreover, we are not confident that the observed positive effect on ICT skills reported by Sharp et al. (2003a) is actually a result of the sport-related extra-curricular activities. The study's data collection and analysis of the ICT skills lowered the overall quality judgement to medium for this outcome.

Figure S2: Forest plot of the impact of extra-curricular learning activity in sports setting on numeracy (compared with no extra-curricular learning activity in sports setting)



Heterogeneity statistic $Q = 0.974$ $df = 1$ $p = 0.324$ $I^2 = 0\%$, $Q^* = 0.974$, Meta-analysis method: random effects model

Figure S3: Forest plot of the impact of extra-curricular learning activity in sports setting on ICT skills (compared with no extra-curricular learning activity in sports setting)

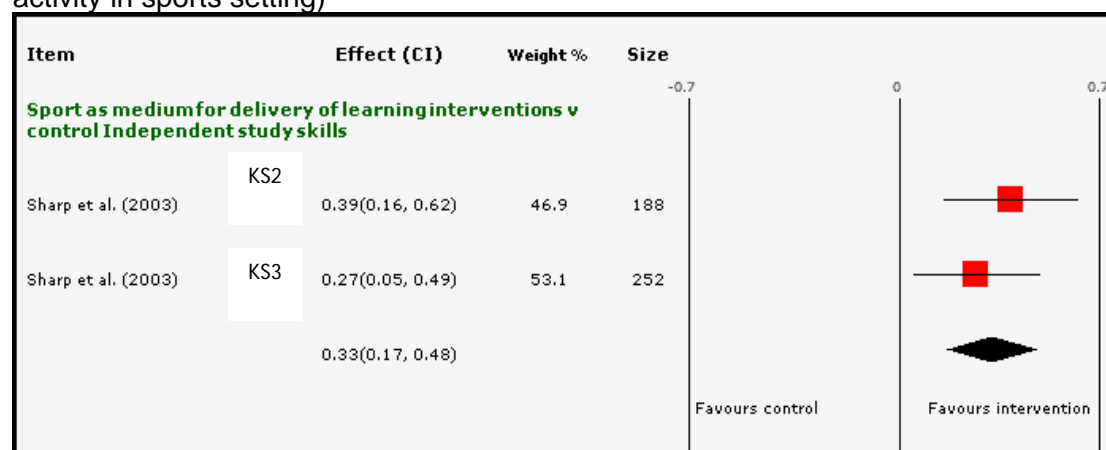


Heterogeneity statistic $Q = 4.67$ $df = 1$ $p = 0.0307$ $I^2 = 78.6\%$ Meta-analysis method: random effects model

Do extra-curricular learning activities linked to organised sport improve transferable skills for underachieving pupils?

This synthesis suggests that students who took part in the extra-curricular activities linked to sport improved their transferable skills (independent study skills) more than a pool of comparable students who did not take part in these extra-curricular activities. The effect sizes (shown in Figure S4) in both Key stages 2 and 3 are positive and exclude the possibility of no effect. The individual effect sizes are sufficiently similar to be combined and the pooled estimate of effect is positive ($g = 0.33$) and the 95% confidence interval excludes a negative effect. However, these outcomes are graded as medium quality and so caution should be used when interpreting these results.

Figure S4: Forest plot of the impact of extra-curricular learning activity in sports setting on independent learning skills (compared with no extra-curricular learning activity in sports setting)



Other study results

There were two other studies that were included in the in-depth review but excluded from the syntheses. The two studies, Anderson (1999) and Powell et al. (2008), were too dissimilar from other included studies to be incorporated into a synthesis. Anderson (1999) did not report comparable outcomes (measuring social competence rather than academic attainment). Powell et al. (2008) examined an intervention and population group that was unlike any other study. A summary of the studies is provided below.

Anderson (1999) investigated the impact of martial arts training on the social competence of primary school boys in the USA. The study found that participation in Taekwondo, compared with no sporting participation, did not have a significant effect (positive or negative) on social competency. However, as the authors point out, the groups of students were not comparable at baseline. The comparison group had higher social competency scores than the karate group in the first place.

Powell et al. (2008) investigated the impact of a massage, yoga and relaxation programme on primary school children in the UK. This intervention, therefore, included elements of non-sporting activity. Entitled the Self Discovery Programme, this intervention was specifically targeted at children with emotional and behavioural difficulties. The study measured the impact of these techniques on a range of behavioural and transferable skills. Powell et al. (2008) report that the Self Discovery Programme led to an increase in student confidence, a reduction in student difficulties and an improvement in student communication. The intervention, however, did not have an observable effect on student concentration or behaviour.

Discussion and implications

Implications for policy and practice

Research evidence does not on its own provide justification for any particular policy decision. Thus, the 'implications' provided here are the review team's interpretation of what we think the balance of evidence indicates is the answer to a particular question.

Young people's participation in organised sport improves their numeracy skills. The pooled estimate of effect ($g = 0.19$) indicates that students' participation

in sporting activities could increase numeracy scores, on average, by 8% above that of non participants (all other things being equal).

Underachieving young people who participate in extra-curricular learning activities linked to organised sport improve their numeracy and transferable skills. These findings apply to both primary and secondary aged school children. The evidence about the effect of this intervention is drawn from Playing for Success and the caveats associated with this study have been noted. It is our judgement, however, that the overall pattern of results suggests that the programme does have beneficial impacts.

1. **Numeracy skills for underachieving young people :** The pooled estimate of effect ($g= 0.80$) indicates that students' participation in extra-curricular learning activities linked to organised sport could increase numeracy scores, on average, by 29% above that of non participants (all other things being equal).
2. **Transferable skills for underachieving young people:** The pooled estimate of effect ($g= 0.33$) indicates that students' participation in extra-curricular learning activities linked to organised sport could increase their performance on transferable skill tests, on average, by between 12% and 16% above that of non participants (all other things being equal).

There is insufficient evidence about the impacts of any other forms of sports participation. However this does not mean that other interventions are ineffective.

Issues to consider

There are three specific issues that need to be taken into account when interpreting the improvements in academic performance and skills that are reported in this review. First, all the studies have compared two groups of young people: one group that participates in organised sport and the other does not participate (comparison group). However, whilst the comparison group may not be participating in the organised sporting activity, this does not mean that this group is not participating in any sport at all. Therefore, the in-depth review can only offer insights into the impacts of organised sport rather than sport per se. Second, this in-depth review evaluates the impact of an intervention (Playing for Success) described as 'extra-curricular learning activities linked to organised sport'. This intervention does NOT only include sporting activities but sport-related activities that are accompanied by learning activities. Therefore, sport is not the primary mechanism for delivering the learning outcome but rather provides a site or setting for the delivery of extra-curricular learning. It is not clear whether, and how, important the sporting context is for achieving the improvements in academic attainment and transferable skills that are reported here. Third, the statistical analysis used to synthesise the data assumes that all of the participants are completely independent of one another. However, this is not the case as individual random allocation is not used in any of the studies. If participants are not independent of one another, the results will be less conclusive than the findings reported above.

Aside from the concerns raised above, there are wider considerations that need to be taken into account when drawing policy or practice implications. These include contextual knowledge about policy and practice in the sporting sector, data about economic costs, knowledge of the desirable and undesirable consequences of (in)action, and awareness of other potential means (aside from those considered in this review) for achieving the same outcomes. Policy or practice discussions would also need to consider the generalisability of the interventions and findings reported here. It is important to note that the interventions have taken place within a particular

context and so the findings may, or may not, be appropriately translated to different contexts. The findings presented here are intended to generate further discussion about the relevance and transferability of the interventions evaluated.

Implications for research

One of the aims of this review was to highlight future directions for research, directed by the gaps that have been identified in the current evidence base.

Secondary research

The systematic map (conducted in stage two) identified research evidence that included quantitative measures of the impact of sporting engagement on learning and/or social outcomes. The map included 268 studies. This in-depth review has only examined a small part of that evidence base (six studies that focused on the impacts of sport on young people's learning outcomes). Future research may wish to explore additional areas of the systematic map. In-depth reviews could be conducted, for example, on the impact of sports participation on different population groups (e.g. the older generation or women only) or for different outcomes (e.g. social outcomes). The in-depth review reported here only synthesised 'high quality' studies. Further in-depth reviews could also incorporate 'low quality' studies into the synthesis to explore the findings reported here.

Future analysis could compare the findings of different systematic review findings. The effect of sport related interventions on young people's learning could be usefully compared with other non-sporting alternatives. Personal development for teachers, for example, has also been found to improve student academic outcomes (Timperley et al. 2007). Whilst effect sizes obtained from different meta-analyses (e.g. sports participation and CPD for teachers) are not directly comparable to the findings reported here, further analysis of could provide further useful information for policy decision making.

Primary research

The in-depth review identified only six 'high quality' studies that examined the impact of young people's participation in organised sport on their learning outcomes. This is a relatively small yield, suggesting that further primary research is required. In order to develop the knowledge base, studies need to use designs that adequately control for bias and have sufficiently large populations to facilitate 'transfer' into policy and practice. With some qualifications, the evaluation of 'Playing For Success' is a model that is worth developing further, with greater consideration given to the use of prospective random allocation rather than retrospective statistical controls. There also appears to be little, if any, cost-effectiveness analysis in this area, so such analysis should be a feature of any new evaluation research that is commissioned.

Strengths and limitations of this systematic review

The systematic map and in-depth reviews/synthesis undertaken as part of 'understanding the drivers, impact and value of engagement with culture and sport' project are, as far as we are aware, the first to apply systematic review techniques to the questions considered. Furthermore, they are the first reviews that have attempted to be comprehensive, systematic and transparent across such a wide area of interest in culture and sport. Because of this, they provide an important resource for the culture and sport communities, not only in their content but also in the development of methods used to investigate the impact of sporting and cultural engagement. These methods can be further applied to many of the questions that remain unanswered in the field.

There are, however, limitations of the review processes that should be taken into account when considering the results that are presented here. There is a detailed presentation and discussion of the limitations of these processes in the Technical Report published alongside this report on the CASE website and only a summary is provided here.

The general approach taken by the project was to provide evidence that could be used within a decision making framework that assessed the relative impacts and values of investments in different kinds of cultural and sporting activities. A different approach may have led to different results. The scope of the review (in terms of what is defined as 'sport') was limited to those sports listed in *Taking Part*. The process of initial searching, whilst comprehensive, was limited to studies in the English language, and furthermore some studies identified in the searches could not be obtained (mainly US dissertations). A limited search of the 'Value of Sport Monitor' was undertaken (unavailable for the duration of the searching), and Sport Information Research Centre, Canada was not searched at all. The use of text mining technology facilitated the initial broad scope of the review but this may have led to relevant studies not being identified in the selection process.

Some limited quality assurance of the search and selection processes was undertaken to assess the extent to which studies may have been missed either in the initial search or by the text mining. We screened reference lists of a recent review of sports participation to identify studies that, based on their title, would appear to have been relevant to this review and thus should have been in our database (see Table S1). We also screened the list of studies provided in the Value of Sports Monitor in March 2010 in the category 'Education and lifelong learning'. Of the twenty three studies listed, twelve were already included in the database; five additional studies were identified for inclusion in the database, none of which met the inclusion criteria for the in-depth review.

The results suggest that our initial search did not identify some relevant studies. A large proportion of the missed studies are US dissertations, which according to the authors of the original reviews, were obtained through personal contacts. The text mining does not appear to have missed any studies. All relevant items identified during the quality assurance exercise were manually entered into the database.

Table S1: Results of the quality assurance exercise

| | Relevant | Missed studies | | | Consequences |
|--------------------|-----------------------------|----------------|-----------------------------|-----------------------|---|
| | Relevant Total ¹ | Missed Total | Missed by search strategies | Missed by text-mining | 'Missed' studies that would be included in the in-depth review ² |
| Long et al. (2009) | 28 | 20 | 20 | 0 | 0 |

1. Based on titles screened against the database inclusion criteria

2. Based on titles screened against the in-depth review inclusion criteria

The selection of studies for the in-depth review was limited to medium to high quality quantitative studies. The restriction to quantitative studies followed the general approach to the project but it could be argued that this is too restrictive a definition of 'impact'. The restriction to high quality studies based on the Maryland Scientific Methods Scale, whilst defensible, may have led to the exclusion of some studies that were in fact 'as good as' those studies that were included, particularly at the cut-off boundary between levels two and three of the scale which are not necessarily as clear as the scale itself might suggest.

The learning impact for young people attending school libraries: an in-depth review

Mark Newman, Jan Tripney, Karen Bird, Irene Kwan, Carol Vigurs, Naira Kalra & Mukdarut Bangpan

Research policy and practice background

Recent government policies have emphasised the value of maximising the public's engagement in culture and sport and demonstrated a commitment to increasing access to these sectors. This is evident in both national and regional strategies (e.g. *Every Child Matters; A Passion for Excellence*) and across government departments (e.g. *Sea Change; London 2012 Olympic Legacy Action Plan; National School Sport Strategy*). Funded bodies are similarly placing great emphasis on increasing and sustaining participation in culture and sport (e.g. English Heritage's *A Lasting Legacy; The Arts Council's Great Art for Everyone*). The current political climate, strengthened by recommendations from the McMaster Review (2008) and recent events (e.g. Liverpool Capital of Culture 2008; preparations for the 2012 Olympics), provides a genuine opportunity for maximising public engagement with culture and sport.

Many claims have been made about the impact of the cultural and sports sectors on a diverse range of outcomes, including education, employment, regeneration, health, social capital, the economy, social inclusion and reducing crime (see for example Arts Council England 2004, Coalter 2005, Maeer 2007, Naylor et al. 2006). However the claims are based on a variety of types of knowledge/evidence. The reviews of such knowledge and evidence, having been prepared for different purposes for different audiences, are not consistent in their approach to gathering, selecting, analysing or quantifying the evidence they present. Thus there does not appear to be a comprehensive, systematic, consistent cross-sectoral analysis of the impacts of culture and sport.

School libraries and young people's learning

The provision of school library services in England varies across local authorities and there would appear to be little systematic information about the different models of school library provision currently in use (Williams, Wavell and Coles, 2001). Visiting libraries and library activities are generally viewed as enjoyable recreational activities by children and young people themselves and by intermediaries such as teachers and parents (Burns Owens Partnership 2009, p.22). Reading, and literacy in particular, are important to academic success, and these form the core of the school library service (Coalter, 2001). However, previous reviews have noted that whilst there is an abundance of anecdotal evidence, extensive 'hard' evidence of impact gathered systematically is often lacking (Williams, Coles and Wavell, 2002). Consequently, these reviews and the data used by them have been argued by their authors to indicate 'the potential' of school libraries under certain conditions (Coalter, 2001).

Authors, funders, and other users of the review

The project was funded by the Culture and Sport Evidence (CASE) programme.¹⁸ The CASE programme comprises the Department for Culture, Media and Sport (DCMS), the Arts Council England (ACE), English Heritage (EH), Museums, Libraries and Archives Council (MLA) and Sport England (SE). The CASE Board is made up of the research directors and/or managers for each organisation, who acted as an Advisory Group.

The project was undertaken jointly by the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London and Matrix Knowledge Group. The EPPI-Centre was responsible for the systematic review work.

Review question

The in-depth review reported here addressed one specific sub-question:

What is the impact of young people's attendance of school libraries on their learning outcomes?

What were the findings of the studies?

Overall descriptive summary of the studies

12 studies examined the impact of young people's attendance of school libraries on their learning. Descriptive summaries for each of the studies are provided in Libraries Appendix L2, p. 125. A full reference list for all the studies cited in this review is given at Libraries Appendix L1, p. 123. All of the studies measured learning outcomes for young people aged between 4 and 16. Ten studies included the secondary school age group, ten primary school age (not mutually exclusive) and in two cases the age of the students was not specified. Four studies examined a mixed sex population and eight studies did not specify the sex of the young people. Ten studies were conducted in the USA, one in Israel and one in Canada. Eight of the studies aimed to measure the impact of better quality library provision on students' academic attainment, that is, the studies compared academic achievement between schools with 'more developed' and 'less developed' library provision. In the other two studies, students were asked about how helpful they perceived school libraries to be. There was one medium quality study; the rest were rated either low or low/medium. This means we should be very cautious about over-interpreting the claims for impact made in these studies.

None of the studies included in this review were theory driven in the sense that they were designed to test a specific hypothesis arising from the operationalisation of a particular theory. All used large-scale survey-type methods and were retrospective rather than prospective. It would appear more that the authors subscribed to a general theory that school libraries improved young people's learning in a number of ways:

- by providing access to a variety of resources for learning
- by providing access to direct instruction
- by working with other educators to design learning strategies
- by providing stimulation/motivation to learn.

¹⁸ <http://www.culture.gov.uk/case/>

Does better quality school library provision improve students' attainment?

All of the studies claim that more developed library provision is associated with better academic attainment. However, the low quality of the studies means that we cannot be confident of the validity of such a claim. Two syntheses were conducted and are reported separately below.

3.2.1 School library provision and academic attainment

In these studies, the nature of the 'more developed' library provision varied between studies and included: longer opening hours; before or after school opening hours; providing trained librarians; more books; provision of classroom libraries as well as central school libraries; changing the role of school librarians; and/or providing improved media equipment for school libraries. In all studies except one, attainment was measured using national or state standardised tests either of a single subject (numeracy/literacy) or a cross-subject average grade score. One study was rated as medium quality and the others as medium low quality.

Box L1: Examples of library provision and attainment studies

Michie and Chaney (2009) examined the effectiveness of the US nationwide programme 'Improved Literacy through School Libraries' (LSL). This programme provided grants to school libraries to increase access and improve facilities. The programme was aimed at schools where a proportion of the student population (20%) were from low-income families. The grants were used by schools to, for example, extend school library opening hours or increase expenditure on books. The programme aimed to improve the literacy skills and academic achievement of young people. To assess the effectiveness of the programme, the authors used surveys and school-level state assessment scores to compare 400 grant-receiving school libraries with 400 non-grant receiving libraries.

Lance et al. (2000a) assessed the impact of Alaska school librarians on academic achievement in 211 public schools (elementary and secondary schools for children aged 6-16 years). The librarians aimed to play an effective instructional role and improve student performance. Using survey data, the authors examined the direct relationship between student performance in literacy and numeracy and library staffing levels, hours of operation, staff activities, usage, technology and cooperation with public libraries. Similar studies were carried out using the same methods in Colorado, Pennsylvania and New Mexico in subsequent years (Lance et al. 2000b, 2000c, 2002).

The medium quality study by Michie and Chaney (2009) measured the effects of a grant, given to schools to improve facilities, on test scores. The change in test scores of schools which had received the grant was compared to the change in test scores of schools which had not received the grant during the same period. A positive association was found between receiving a grant and improved reading outcomes; in the case of secondary schools, it was statistically significant. The regression analysis included statistical controls for other contemporaneous school reforms, and when these were included in the regression models, the difference remained positive but no longer statistically significant ($p=0.052$)

The other seven studies compared academic attainment between schools with more developed or less developed library provision at one point in time, i.e., they did not measure a change in attainment scores linked to an improvement in provision. The

studies used similar methods to survey schools about library provision and combined this with secondary data analysis of databases of state/regional test scores. The results from the studies are presented in slightly different ways using multiple and varying measures of the independent variable (library provision). Whichever measure of library provision was used, the results remained the same (in terms of the direction of effect).

Table L1: Results Academic attainment in schools with more developed or less developed library provision.

| Study | Primary school | Middle school | Secondary School |
|-------------------------------------|----------------|---------------|------------------|
| Baughman (2000) | + | + | + |
| Lance et al. (2000a) (Alaska) | + | + | + |
| Lance et al. (2000b) (Colorado) | + | Not in study | + |
| Lance et al. (2000c) (Pennsylvania) | + | + | + |
| Lance et al. (2002) (New Mexico) | + | + | + |
| Mitchie and Chaney (2009) | + | Not in study | + |
| Queens University (2006) | ¹ | Not in study | Not in study |
| Rodney et al. (2003) | + | + | + |
| Shoham (2000) | +/- | Not in study | Not in study |
| Smith (2001) | + | + | + |

1. Statistical significance not reported

* Authors report statistically significant

The main results from the studies are shown in Table L1. The + sign in the table indicates that the result favoured the schools with greater levels of provision. The study by Shoham (2000) presents results for three groups: schools with classroom libraries only, schools with central libraries only, and schools with both classroom and central libraries. It is not clear which of these can be viewed as the 'more developed' form of provision and the results presented by the authors favour different groups depending on the measure of reading attainment used. All of the studies with the exception of Shoham (2000) use regression models that contain other variables. The models vary in sophistication in terms of the number of variables or factors that are controlled for in the regression coefficient used in the results table. As a minimum all use some form of control for differences in socio-economic status between schools.

The results from all of the studies included in Table L1 suggest that more developed library provision is associated with better academic attainment. However, the design of these studies means that we cannot be confident about the direction of the cause and effect relationship. Although these studies 'control for' differences between schools in socio-economic status, this measure relates only to the point at which the outcome was measured after the intervention. It does not take into account the issue of any baseline differences that there may have been between schools in reading achievement. This is achieved by measuring a change score and controlling for differences in other factors that may affect the outcome. Where this kind of analysis

was undertaken in the study by Michie and Chaney (2009), the greater improvement in test scores seen in secondary schools was no longer statistically significant. However, this does not necessarily mean that it was unimportant, as the authors state that the average level of reading score improvement amongst grantees was 2.2% higher than for non-grantees, even when the effect of other educational reforms was taken into account.

The perceived 'helpfulness' of school libraries in improving attainment

The two other studies that investigated the impact of school libraries (Todd and Kuhlthau 2004, Smith 2006) used a different approach from those discussed above. They investigated the perceptions of students about the 'helps' (the term used by the authors) that they received from school libraries. The studies used the same methods and instruments, which comprised a survey in which students rated a particular 'help' on a 5-point Likert scale. The authors reported results for each individual 'help' (48 listed). The 'helps' were also put into groups which the authors argued reflected distinct different types of impacts.

Box L2: Example study 'helpfulness' of school libraries

Smith (2006) explored the ways and extent to which students benefit from school libraries. Using survey data from 3,957 students (aged 6-18 years) and 107 teachers from 51 schools in Wisconsin, the author presented students' and teachers' report on the helpfulness of school libraries. Results showed that over two-thirds of students and teachers rated the school library services 'helpful' and 'very helpful' in getting and using Information for school work, in using computer technology, in increasing reading interest, in improving knowledge and academic achievement, and in helping students to become independent learners.

Those which we considered most directly linked to learning and attainment outcomes are shown in Table L3.2. It is difficult to interpret the results of these studies. How, for example, should a mean score of 2.8 be interpreted on a scale that runs from '1 very helpful' to 5 'not at all helpful'. The authors argue that the results show that on average the students found school libraries helpful or had an impact to some degree in all these areas.

Table L2: Mean scores* for student ratings of learning/attainment help provided by school libraries

| Group | Smith (2006) | Todd and Kuhlthau (2004) |
|----------------------|---------------------|---------------------------------|
| Knowledge | 2.8 | 2.1 |
| Reading | 2.8 | 1.9 |
| Independence | 3.0 | 1.8 |
| Academic achievement | 2.8 | 2.0 |

* Scale for calculating means ranged from '1 - very helpful' to '5 - not at all helpful'

Discussion and implications

Implications for policy and practice

Research evidence does not on its own provide justification for any particular policy decision. Thus the 'implications' provided here are the review teams' interpretation of

what we think the balance of evidence indicates is the answer to a particular question.

Based on the findings of this review, the interpretation framework suggests that there is promising evidence that improving the quality of school libraries improves academic attainment. There is one medium quality study in which the direction of effect suggests this (although the results are not statistically significant) and the preponderance of other evidence also supports this interpretation. Further high quality studies could be carried out to verify this finding, particularly in the UK context.

There are of course many issues that need to be taken into account when considering what if any actions these interpretations suggest. These include knowledge of policy and practice in the field, knowledge about economic costs, knowledge about values and beliefs, knowledge about other desirable and undesirable consequences of (in) action, and knowledge about other potential means for achieving the same ends which have not been investigated in this review. Such discussion would also need to consider the generalisability of any interventions and study results from the context in which they were generated into the context in which they may be used. Discussion of this type is not the role of this review and as such the findings presented here form the beginning rather than the end of any investigation and analysis in this area.

Implications for research

One of the aims of this review was to highlight potential research directions suggested by the current evidence base in this area.

Secondary research

This review has looked at only a small part of the evidence base for the impact of school library attendance and/or provision. Future research may wish to explore additional areas of the systematic map, for example, by undertaking a synthesis of the impact of school library attendance and/or provision on other groups and for other outcomes. Further reviews could be undertaken to examine the impact of other types of libraries (specialist and general) on young people's learning. We have already noted that the database of studies from which this review was drawn is unlikely to include studies about the impacts of early literacy/reading schemes in which libraries are involved. New searches could be carried out to identify studies that investigate the impact of such interventions to add these to the database.

Primary research

Even with the caveat that this review has explored only a small part of the evidence base for the impact of school library attendance and/or provision, it is clear that the yield of quantitative studies in this area is comparatively small and those defined as high quality even smaller. This was also noted in previous reviews of the impact of library provision (Coalter 2001; Williams, Wavell and Coles 2001; Williams, Coles and Wavell 2002) and the results of this review suggest that with regard to high quality quantitative evaluative evidence, this position has not improved greatly in the years since these earlier reviews were carried out. Whilst this is not unusual in most areas of social policy, commissioners of research should consider commissioning further research in this area. This has already been anticipated by MLA, whose action plan for libraries states that MLA will invest in impact research and publish the results to show the impact libraries have on reading, learning and community engagement (MLA 2008).

However, any commissioned research needs to be designed such that it can develop the existing knowledge base. Studies would need to use designs that control

adequately for bias and thus provide more confidence in making causal claims. Studies also need to be sufficiently large scale to facilitate 'transfer' into policy and practice. There also appears to be little if any cost-effectiveness analysis of school library provision, and this could also be a feature of any new evaluation research that is commissioned.

Strengths and limitations of this systematic review

The systematic review and in-depth reviews/synthesis undertaken as part of the 'Understanding the drivers, impact and value of engagement in culture and sport' project are, as far as we are aware, the first to apply systematic review technology to the questions considered (although reviews by Williams, Wavell and Coles (2001, 2002) claim to have used some of the aspects of the approach). Furthermore, they are the first reviews that have attempted to be comprehensive, systematic and transparent across such a wide field of interest in culture and sport. Because of this, they provide an important resource for the culture and sport communities, not only in their content but also in the development of methods to use this approach to continue to investigate many of the questions that remain unanswered in the field.

This review represents only the first step in an ongoing process of building knowledge and understanding about the impact of library attendance and/or provision. The in-depth review addressed only a very small part of the agenda of interest, but importantly, the systematic and comprehensive approach used means that it will be possible to utilise the database of studies to begin to address some of the other questions of interest in subsequent reviews. However there are limitations to the review processes that should be taken into account when considering the results that are presented here. There is a detailed presentation and discussion of the limitations of these processes in the review technical report (Tripney et al. 2010), and only a summary is provided here.

The general approach taken by the project was to provide evidence that could be used within a decision-making framework that assessed the relative impacts and values of investments in different kinds of cultural and sporting activities. A different approach may have led to different results. The process of initial searching, whilst comprehensive, was limited to studies published in the English language and furthermore some studies identified in the searches could not be obtained (mainly US dissertations). The use of text mining technology facilitated the initial broad scope of the review but may have led to relevant studies not being identified in the selection process.

Some limited quality assurance of the search and selection processes was undertaken to assess the extent to which studies may have been missed, either in the initial search or by the text mining. We looked at the reference lists of a recent review on the impact of libraries to identify studies that, based on their title, would appear to have been relevant to this review and thus should have been in our database. The results of this exercise are shown in Table L3.

Table L3: Results of the quality assurance exercise

| | Relevant | Missed studies | | | Consequences |
|--------------------------------|----------------|----------------|-----------------------------|-----------------------|--|
| | Relevant Total | Missed Total | Missed by search strategies | Missed by text mining | Missed studies that would be included in the in-depth review |
| Burns Owens Partnership (2009) | 18 | 8 | 8 | 0 | 0 |

The results suggest that the initial search failed to identify some relevant studies. The text mining does not appear to have missed any studies. However, none of the missed studies would have been eligible for inclusion in this review. All relevant items identified during the quality assurance exercise were manually entered into the database.

The learning impacts for young people attending Museums, Galleries and Heritage sites: an-in depth review

Mark Newman, Karen Bird, Jan Tripney, Irene Kwan, Carol Vigurs, Naira Kalra, Mukdarut Bangpan

Research, policy and practice background

Cross Sector

Recent government policies have emphasised the value of maximising the public's engagement in culture and sport and demonstrated a commitment to increasing access to these sectors. This is evident in both national and regional strategies (e.g. Every Child Matters; A Passion for Excellence) and across government departments (e.g. Sea Change; London 2012 Olympic Legacy Action Plan; National School Sport Strategy). Funded bodies are similarly placing great emphasis on increasing and sustaining participation in culture and sport (e.g. English Heritage's A Lasting Legacy; The Arts Council's Great Art for Everyone). The current political climate, strengthened by recommendations from the McMaster Review (2008) and recent events (e.g. Liverpool Capital of Culture 2008; preparations for the 2012 Olympics), provides a genuine opportunity for maximising public engagement with culture and sport.

Research on the impacts of the cultural and sporting sectors is extensive and wide ranging. Primary and secondary studies have claimed a diverse range of impacts of these sectors on education, employment, regeneration, health, social capital, the economy, social inclusion and crime (for literature reviews see, for example, Arts Council England 2004, Coalter 2005, Maeer 2007, Naylor et al 2006). This research has provided the basis for making knowledge claims about the values and effects of public engagement in sport and culture. These claims, however, are based on different types of knowledge, reviews of such knowledge, and research evidence that has been prepared for differing purposes. There has been no consistent approach used to gather, selecting, analysing or quantifying the evidence to support claims of impact. Whilst there would appear to be an evidence base it is inconsistent fragmented and difficult to interrogate and compare.

Policy: Museums, Galleries and Heritage

During the 1990s, a number of factors prompted the museums sector to create audience development strategies and widen access to their collections (Tlili et al. 2007). Local museums were pioneers in widening access, with documented success of attracting non traditional audiences (Hooper-Greenhill, 1994). Towards the end of the 1990s, these strategies converged with the government's social inclusion agenda, which was formally extended to museums galleries and heritage.

Government policy identified widening access as a top priority, requiring the sector to deliver and monitor interventions to increase participation (DCMS, 1999). Building on this priority, an expanded social role for museums and galleries was established in Centres for Social Change (DCMS 2000a) and recognised in the heritage sector in the Power of Place (English Heritage 2000). These policy documents recognised the potential of the sector as a vehicle for social change, and the need for social inclusion to be mainstreamed as a priority in museums and galleries.

Research: Museums Galleries and Heritage

Studies of attendance began to emerge in the 1990s, with research examining the views of underrepresented groups and barriers to engagement (Desai and Thomas 1998; MORI 1999). Studies also began to interrogate the value and impact of cultural institutions. Early studies considered the impacts on local, regional and national economies (Travers 2006). Subsequently, a range of other impacts have been measured. Studies examining social exclusion, neighbourhood renewal and community cohesion provide an established evidence base. In comparison, there would appear to be less research on the effects of museums galleries and heritage on health, regeneration and cultural diversity. It is argued that the strongest evidence of social impacts pertains to individual development and learning, specifically in 'life skills' (Burns Owens Partnership, 2005).

Museum, gallery and/or heritage site attendance and young people's learning

The nature and extent of the educational role of the museum grew dramatically in the 1990s. From providing specific activities for particular population groups (such as schoolchildren), the museum's education role grew to include exhibitions, displays and workshops (Hooper-Greenhill 1998). For policy makers, education and lifelong learning has become a central part of the socially inclusive museum and gallery. *The Learning Power of Museums: A Vision for Museum Education* (DCMS 2000b) sets out the government's vision for museums and galleries. More recently '*The Government's Statement on the Historic Environment in England 2010* (DCMS 2010) promoted the importance of the historic environment as a focus for community identity and learning. The sector can support young people's learning outcomes by enhancing the delivery of the national curriculum, bringing classroom teaching to life, and fostering key skills and creative abilities. Government objectives provided the impetus for the sector to measure its impacts on educational and other outcomes (Wavell et al. 2002). A full list of references for this review is given at Museums Appendix M1, p. 137.

Authors, funders, and other users of the review

The project was funded by the Culture and Sport Evidence (CASE) programme¹⁹. The CASE programme comprises the Department for Culture, Media and Sport (DCMS), the Arts Council England (ACE), English Heritage (EH), Museums, Libraries and Archives Council (MLA) and Sport England (SE). The CASE Board is made up of the research directors and/or managers for each organisation who acted as an Advisory Group.

The project was undertaken jointly by the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London and Matrix Knowledge Group. The EPPI-Centre was responsible for the systematic review work.

¹⁹ <http://www.culture.gov.uk/case/>

Review question

The in-depth review reported here addressed one specific question:

What is the impact of attending museum, gallery and/or heritage on young people's learning outcomes?

What were the findings of the studies?

Overall descriptive summary of the studies

Eleven studies investigated the impact of museum, gallery and/or heritage site attendance on young people's learning outcomes. In all of the studies, the students in the intervention group (or all students in the case of studies where there was not a control or comparison group) attended either a museum, art gallery and/or some kind of heritage site. These labels are not used consistently between studies and in some studies the students attended different types of cultural institution. Descriptive summaries for each of the studies are provided in Museums Appendix M2, p. 140. This set of studies was split into two groups. The first group of studies (n=5) evaluated the impact of museum, gallery and/or heritage site attendance that included additional learning support from professionals or supplementary materials. Three of these studies were conducted in the US (Melber 2003; Randi Korn Associates 2007; Kanevsky et al. 2008) and two in the UK (AEA Consulting 2005; Stanley 2008)

The second group of studies (n=6) evaluated the impact of museum, gallery and/or heritage site attendance only (the studies do not report that further learning support is provided for young people). Young people's museum, gallery and/or heritage site attendance took place through organised school visits to galleries or museums. All of these studies were undertaken in the UK (Hooper-Greenhill et al. 2004a, 2004b, 2006, 2007; Watson et al. 2007; West 2008). All of the studies measured 'learning' outcomes for mixed sex populations aged 4-16.

Box M1: examples of studies of museum, gallery and/or heritage site attendance, with additional learning support

Randi Korn Associates (2007). This study examined the impact of the LTA (Learning Through Art program) on students and teachers in the New York City public school system from the 2004-06. The LTA program placed practising artists in New York City Schools to work with students (aged 6-10 years) and classroom teachers on curriculum-based art projects. The 20-week LTA program aimed to improve scores in the English Language Arts Test (ELA) by promoting positive attitudes toward school, art, and art museums in third graders (n=605). In a quasi-experimental design using questionnaires, interviews, observations and case studies, the authors assessed the effectiveness of the LTA program by comparing students in the treatment group who received the LTA program, with students in a non treatment group who did not receive the LTA program.

Kanevsky et al (2008) assessed the effectiveness of the School in the Park (SITP) intervention, intended to nurture the resilience of inner-city students in third to fifth grades (aged 6-10 years) who were economically and academically challenged. SITP was an enriched, museum-based intervention, in which students worked with museum educators daily for 2-3 months. Using an experimental design, the authors compared the academic

resilience and personal development of SITP participants with those of non-participants over two years.

Melber (2003) investigated the impact of a specially designed museum science program on academically gifted elementary students. The programme ran mainly in school-based extra-curricular clubs (i.e. those taking place in schools in out-of-school hours) with one session at a museum. The sessions incorporated participation activities that were similar to the working activities of museum scientists and incorporated actual museum specimens and artefacts.

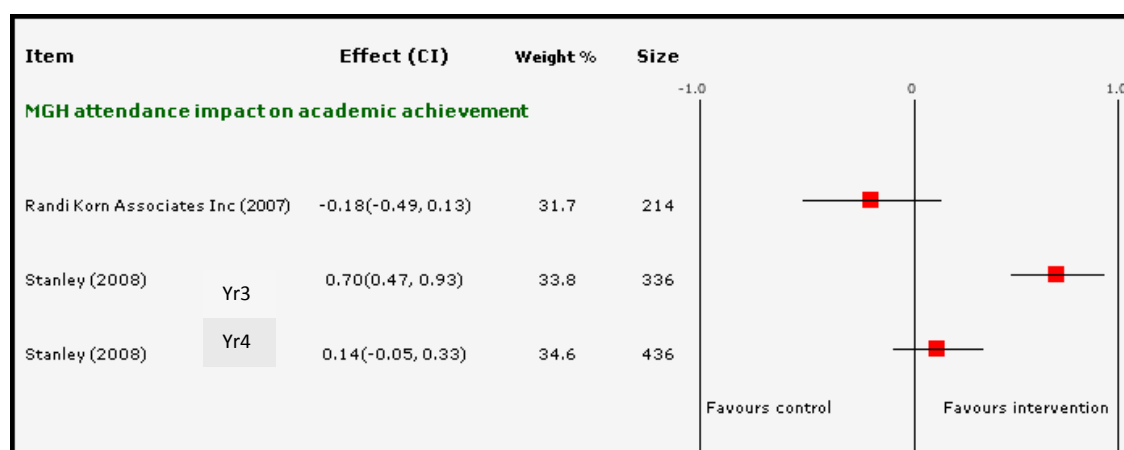
Does museum, gallery and/or heritage site attendance with supplementary learning support improve the learning outcomes of young people?

This group of studies investigated whether attending museums, galleries and/or heritage sites in conjunction with or as part of a programme that was run jointly by schools and cultural institutions staffed both by school teachers and professional staff from the cultural sector, improved student learning outcomes. Studies investigated a range of learning outcomes. A synthesis was carried out for each outcome type.

Does museum, gallery and/or heritage site attendance with supplementary learning support improve the academic attainment of young people?

Two studies measured the impact of museum, gallery and/or heritage site attendance on academic attainment (Randi Korn Associates 2007; Stanley 2008). These studies used statewide or national tests to measure the impact on literacy. The outcome measures for both studies are plotted in Figure M1 below. This forest plot includes two effect sizes for Stanley (2008), one for year 3 (aged 7/8) and one for year 4 (aged 8/9).

Figure M1: Forest plot of the impact of museum, gallery and/or heritage site attendance (with supplementary support) on young peoples' academic achievement



Heterogeneity statistic $Q = 24.1$ $df = 2$ $p = 5.97E-06$ $I^2 = 91.7\%$, $Q^* = 2.26$ random effects model

As illustrated by the forest plot, the individual studies report very different findings. The effect size calculated from Randi Korn Associates(2007) indicates that the students who did not participate in the 'museum, gallery and/or heritage site attendance activities' got higher test scores than the students who did, whereas the effect sizes calculated from Stanley (2008) indicate the opposite. The heterogeneity statistics confirm that the individual study results are too dissimilar to combine and

therefore no pooled estimate of effect is presented. It is not clear why the studies should find opposite effects. Both studies are low quality which that suggests that biases arising from study limitations (e.g. unequal groups) are a possible factor in either or both.

Does museum, gallery and/or heritage site attendance with supplementary learning support improve the 'attitude toward school' of young people?

This outcome was measured by three studies (Kanevsky et al. 2008; Randi Korn Associates 2007; AEA Consulting 2005). Two of these included data that were suitable for meta-analysis and the results of these are presented on in Figure M2. This includes two effect sizes from Kanevsky et al (2008), one for students at grades 3/4 and one for those at grades 4/5. The students in the Randi Korn Associates study were in grade 3. The results for all three of the outcomes are inconclusive, as is the pooled estimate of effect. The remaining study (AEA, 2005) surveyed students who had participated in a Museums outreach programme and 75% of the students of those who responded said that the programme had increased their desire to learn.

Does museum, gallery and/or heritage site attendance with supplementary learning support improve the learning capabilities of young people?

Two studies (Kanevsky et al. 2008; Randi Korn Associates, 2007) measured a phenomenon that could be described as students' *capability to learn*. The two studies use different terms to describe this concept ('academic self efficacy' and 'artistic process' respectively) but used data collection tools that include similar items. The effect sizes calculated from these studies are presented in **Figure M3**. This includes two outcome measures for Kanevsky et al. (2008), one for students at grades 3/4 and the other for those at grades 4/5. The results in all three of the outcomes are inconclusive, as is the pooled estimate of effect.

Figure M2: Forest plot of the impact of museum, gallery and/or heritage site attendance (with supplementary support) on young peoples' attitude towards school

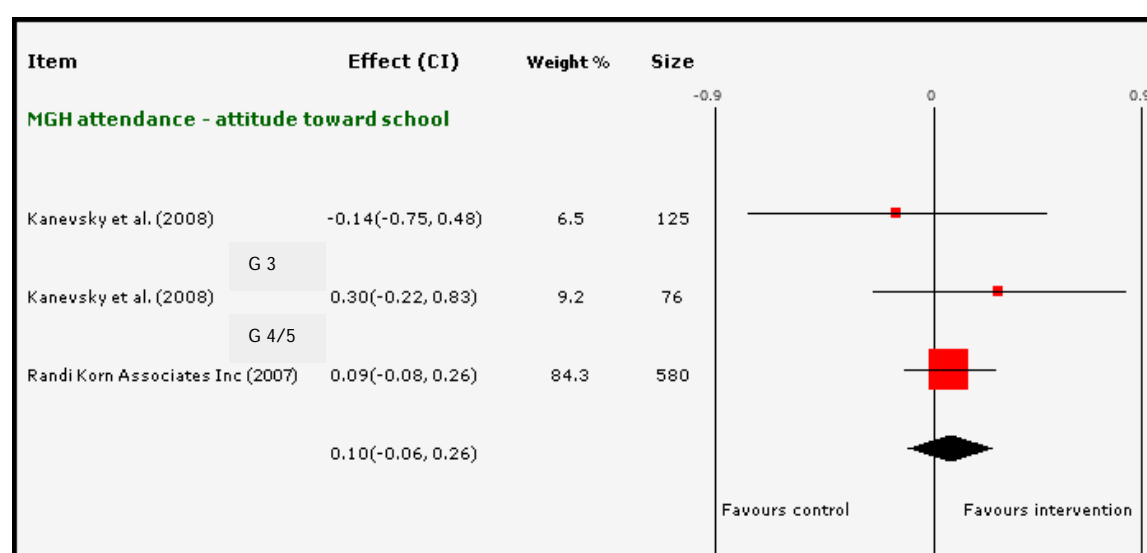
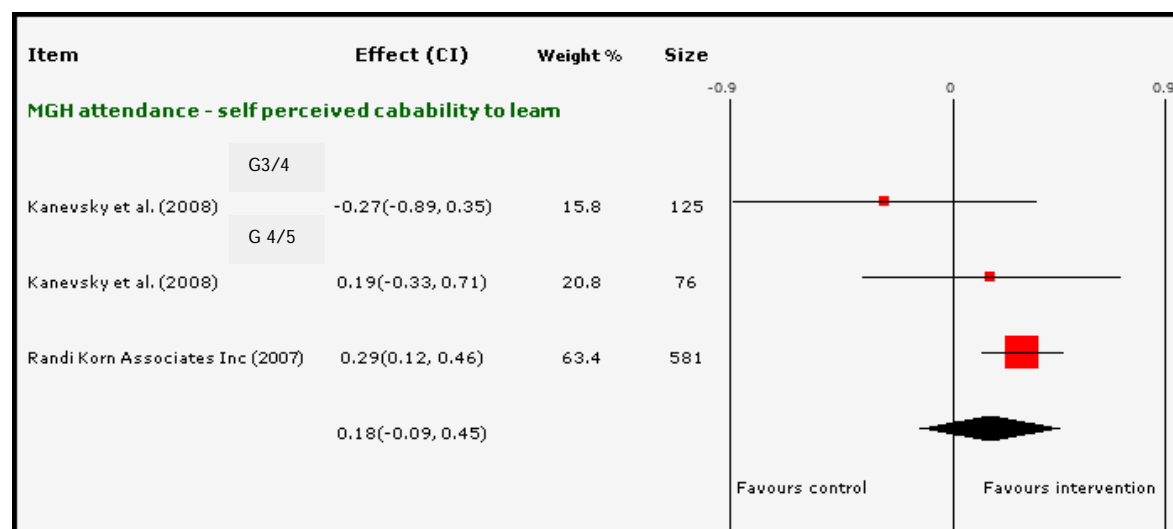


Figure M3: Forest plot: The impact of museum & gallery attendance outreach on young people's self-perceived capacity to learn



Heterogeneity statistic $Q = 2.97$ $df = 2$ $p = 0.227$ $I^2 = 32.6\%$, random effects model

Other outcomes

The studies also reported unique outcomes that could not be combined. Melber (2003) measured the impact of museum attendance on student's views of careers in science. The study found that participants had a more positive view of such careers and fewer stereotypical views after attending a science museum than before. Randi Korn Associates (2007) measured student attitudes towards studying art and visiting art museums. They report that participants in the intervention had, on average, more positive attitudes towards art than non participants. AEA Consulting (2005) measured student's perceptions of impacts on their confidence and creativity. The study found that 100percent of students reported an increase in their confidence, and 83 percent claimed that the programme had developed their creativity.

Does museum, gallery and/or heritage site attendance (organised via school) improve the learning capabilities of young people?

Six studies examined the impact of young people's museum, gallery and/or heritage site attendance (with their school) on learning outcomes. All of these studies measured perceived impacts only. They were conducted in the UK and all used what might be described as a single-group post-intervention only design in which students and /or their teachers were asked by means of questionnaires to state what or how the visit had impacted on their learning. Two studies (Watson et al. 2007; West 2008) also reported teachers' perceptions of impact of the visit on students' academic grades.

It is difficult to know how to synthesise these studies as each one was graded as being of low quality with respect to the review question, the answers are self-perceived responses to questionnaire items with no reported validity or reliability, response rates were less than 20% in all cases and none had any statistical analysis (just presenting frequencies). All of the studies measured one or more of the following:

- Enjoyment,

- Interest of the visit
- Impact on subject understanding
- Understanding of what they had seen at the museum/gallery/heritage site
- Impact on motivation for schoolwork
- Whether the learning was useful for things other than school work
- Whether the visit was exciting

In each case the results were overwhelmingly positive with usually between 70 and 80 percent of respondents giving positive ratings. Positive ratings were almost always higher in younger than older students and amongst girls.

With regard to the impact on attainment in the study by Watson et al. (2007) teachers perceived that academic assignment grades had increased in more than 60 percent of pupils who had attended the museums and fallen in about 20 percent. The results with regard to academic attainment in the study by West (2008) are not clearly reported but the author seems to suggest that there were improvements in some aspects of academic performance as a result of the museum visits.

Box M2: Example Museum, gallery and/or heritage site attendance

Watson et al. (2007) examined the impact that museums in the East of England region had on the attainment of 762 secondary-age pupils (aged 11-18 years). Data of students' academic performance were obtained from school records, based on the students completing an assessed piece of work after one museum visit with their teachers. The students also completed a questionnaire survey about their learning experiences after their museum visit.

Discussion and Implications

Implications for policy and practice

Research evidence does not on its own provide justification for any particular policy decision. Thus the 'implications' provided here are simply the review teams interpretation of what we think the balance of evidence indicates is the answer to a particular question. Using the interpretation framework the results would suggest that there is 'promising evidence' that museum, gallery and/or heritage site attendance with supplementary learning may improve students' 'attitudes towards school' and 'self confidence in their learning abilities'. However this is based on the results of only two studies and only one of them reaches the acceptable 'medium' quality threshold set by the CASE programme. There is insufficient evidence about the impact of museum, gallery and/or heritage site attendance with supplementary learning alone on other learning outcomes or about the impact of museum, gallery and/or heritage site attendance alone.

There are of course many issues that need to be considered when determining what if any actions these interpretations suggest, including knowledge of policy and practice in the field, knowledge about economic costs, knowledge about values and beliefs, knowledge about other desirable and undesirable consequences of (in)action, and knowledge about other potential means of achieving the same ends which have not been investigated in this review. Such discussion would also need to consider the generalisability of any interventions and study results from the context in which they were generated into the context in which they may be used. Discussion

of this type is not the role of this review and as such the findings presented here form the beginning rather than the end of any investigation and analysis in this area.

Implications for research

One of the aims of this review was to highlight potential research directions suggested by the current evidence base in this area.

Secondary research:

This review has looked at only a small part of the evidence base for the impact of museum, gallery and/or heritage site attendance. Future research may wish to explore additional areas of the systematic map, for example, by undertaking a synthesis of the impact of the museum, gallery and/or heritage site attendance on other groups and other outcomes.

Primary research:

Even with the caveat that this review has explored only a small part of the evidence base for the impact of museum, gallery and/or heritage site attendance, it is clear that the yield of quantitative studies in this area is comparatively small, and those defined as high quality even smaller. Whilst this is not unusual in most areas of social policy, commissioners of research should consider commissioning further research in this area. Interventions in the 'attendance plus supplementary learning' category could be a priority for future evaluative research efforts.

However, any commissioned research needs to be designed such that it can develop the existing knowledge base. Studies would need to use designs that control adequately for bias and that were of a sufficiently large scale to facilitate 'transfer' into policy and practice. There also appears to be little if any cost-benefit analysis in this area and this should be a feature of any new evaluation research that is commissioned.

Strengths and limitations of this systematic review

The systematic review and in-depth reviews/synthesis undertaken as part of the 'Understanding the drivers, impact and value of engagement in culture and sport' project are, as far as we are aware, the first to apply systematic review technology to the questions considered. Furthermore, they are the first reviews that have attempted to be comprehensive, systematic and transparent across such a wide field of interest in culture and sport. Because of this, they provide an important resource for the culture and sport communities, not only in their content but also in the development of methods for continuing to investigate the many questions that remain unanswered in the field.

This review represents only the first step in an ongoing process of building knowledge and understanding about the impact of museum, gallery and/or heritage site attendance. The in-depth review addressed only a very small part of the agenda of interest, but importantly, the systematic and comprehensive approach used means that it will be possible to utilise the database of studies to begin to address some of the other questions of interest in subsequent reviews. However, there are limitations to the review processes that should be taken into account when considering the results that are presented here. There is a detailed presentation and discussion of the limitations of these processes in the review technical report which is published alongside this report on the CASE website (Tripney et al 2010) and only a summary is provided here.

The general approach taken by the project was to provide evidence that could be used within a decision-making framework that assessed the relative impacts and values of investments in different cultural and sporting activities. A different approach may have lead to different results. The process of initial searching whilst comprehensive was limited studies in the English language and furthermore, some studies identified in the searches could not be obtained (mainly US dissertations). The use of text-mining technology facilitated the initial broad scope of the review but may have led to relevant studies not being identified in the selection process.

Some limited quality assurance of the search and selection processes was undertaken to assess the extent to which studies may have been 'missed either in the initial search or by the text mining. We looked at the reference lists of a recent review on the impact of museums to identify studies that, based on their title, would appear to have been relevant to this review and thus should have been in our database. The results of this exercise are shown in Table M5 below.

Table M5 Results of the quality assurance exercise

| | Relevant | Missed studies | | | Consequences |
|---------------|-------------------|-----------------------|-----------------------------------|------------------------------|--|
| | Relevant Total | Missed Total | Missed by search strategies | Missed by text- mining | 'Missed' studies that would be included in the in- depth review |
| BOP (2005) | 15 | 1 | 1 | 0 | 0 |

The results suggest that the initial search only missed one relevant study, which would not have been included in the in-depth review. This was subsequently entered manually into the database.

Appendices

The appendices are split into two areas:

In-depth review specific areas, covering the 4 reviews:

- Arts in-depth review
- Sports in-depth review
- Libraries in-depth review
- Museums, galleries and heritage in-depth review

General appendices covering the background approach and methods

Arts in-depth review appendices

Arts in-depth review: Appendix A1

References

References for studies in the in-depth review

*Four studies included in the in-depth review were excluded from the synthesis following the process of data extraction and quality appraisal as their overall weight of evidence was judged to be low or low/medium.

Linked reports (i.e., additional publications relating to the following studies) are detailed in section 5.2.

Bigelow IL (1997) Assessing the relationship between participation in the performing arts in schools and communication apprehension. *Dissertation Abstracts International Section A: Humanities and Social Sciences* M1–58: 0714.

Bilhartz TD, Bruhn RA, Olson JE (1999) The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology* 20(4): 615–636.

*Bolduc J (2009) Effects of a music programme on kindergartners' phonological awareness skills. *International Journal of Music Education* 27: 37–47.

Catterall JS (2007) Enhancing peer conflict resolution skills through drama: an experimental study. *Research in Drama Education* 12(2): 163–178.

Catterall JS, Peppler KA (2007) Learning in the visual arts and the worldviews of young children. *Cambridge Journal of Education* 37(4): 543–560.

Costa-Giomi E (2004) Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music* 32: 139–152.

Fleming M, Merrell C, Tymms P (2004) Impact of drama on pupils' language, mathematics, and attitude in two primary schools. *Research in Drama Education* 9(2): 177–197.

Freeman GD (2001) Effects of creative drama activities on third and fourth grade children. *Dissertation Abstracts International Section A: Humanities and Social Sciences* M1–61: 3470.

Gromko JE (2005) The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research in Music Education* 53(3): 199–209.

Gromko JE, Poorman AS (1998) The effect of music training on preschoolers' spatial-temporal task performance. *Journal of Research in Music Education* 46: 173–181.

Hui A, Lau S (2006) Drama education: a touch of the creative mind and communicative-expressive ability of elementary school children in Hong Kong. *Thinking Skills and Creativity* 1(1): 34–40.

Kendall L, Morrison J, Yeshanew T, Sharp C (2008) *The longer-term impact of creative partnerships on the attainment of young people: results from 2005 and 2006. Final report.* Slough: National Foundation for Educational Research.

*Kennedy R, Scott A (2005) A pilot study: the effects of music therapy interventions on middle school students' ESL skills. *Journal of Music Therapy* 42(4): 244–261.

Kim J, Wigram T, Gold C (2008) The effects of improvisational music therapy on joint attention behaviors in autistic children: a randomized controlled study. *Journal of Autism and Developmental Disorders* 38(9): 1758–1766.

Krahe B, Knappert L (2009) A group-randomized evaluation of a theatre-based sexual abuse prevention programme for primary school children in Germany. *Journal of Community and Applied Psychology* 19(4): 321–329.

McMahon S, Rose D, Parks M (2003) Basic reading through dance program: the impact on first-grade students' basic reading skills. *Education Review* 27(1): 104-125.

Merrell R (2005) The impact of a drama intervention program on the response of the bystander to bullying situations. *Dissertation Abstracts International Section A: Humanities and Social Sciences* M1–65: 2503.

*O'Gara P (2008) To be or have not been: learning language tenses through drama. *Issues in Educational Research* 18(2): 156-166.

Orsmond GI, Miller LK (1999) Cognitive, musical, and environmental correlates of early music instruction. *Psychology of Music* 27: 18-37.

Piro JM, 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.

Rauscher FH, Shaw GL, Levine LJ, Wright EL, Dennis WR, Newcomb RL (1997) Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurological Research* 19: 2-8.

Rauscher FH, Zupan MA (2000) Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: a field experiment. *Early Childhood Research Quarterly* 15(2): 215-228.

Register D (2004) The effects of live music groups versus an educational children's television program on the emergent literacy of young children. *Journal of Music Therapy* 41(1): 2-27.

Schellenberg E (2004) Music lessons enhance IQ. *Psychological Science* 15(8): 511-514.

*Schunk HA (1999) The effect of singing paired with signing on receptive vocabulary skills of elementary ESL students. *Journal of Music Therapy* 36(2): 110-124.

Standley J, Hughes J (1997) Evaluation of an early intervention music curriculum for enhancing pre-reading writing skills. *Music Therapy Perspectives* 15: 79-86.

von Rossberg-Gempton IE, Dickinson J, Poole G (1999) The potential for enhancing psychomotor and cognitive functioning in frail seniors and young children through creative dance. *Journal of Human Movement Studies* 37(5): 235-260.

Wright R, John L, Ellenbogen S, Offord DR, Duku EK, Rowe W (2006) Effect of a structured arts program on the psychosocial functioning of youth from low-income communities: findings from a Canadian longitudinal study. *Journal of Early Adolescence* 26(2): 186-205.

Linked reports (additional publications related to studies in the in-depth review)

Costa-Giomi E (1999) The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education* 47(3):198-212.

Eames A, Benton T, Sharp C, Kendall L (2006) *The impact of creative partnerships on the attainment of young people: final report*. Slough: National Foundation for Educational Research.

- Kendall L, Morrison J, Sharp C, Yeshanew T (2008) *The impact of creative partnerships on pupil behaviour*. Slough: National Foundation for Educational Research.
- Merrell C, Tymms P (2002) *The National Theatre Transformation Project: final evaluation report*. Durham: University of Durham.
- Register D (2003) *The effects of live music groups versus an educational children's television program on the emergent literacy of young children*. Tallahassee, FL: Florida State University.
- Sharp C, Pye D, Blackmore J, Brown E, Eames A, Easton C, Filmer-Sankey C, Tabary A, Whitby K, Wilson R, Benton T (2006a) *National Evaluation of Creative Partnerships: final report*. Slough: National Foundation for Educational Research.
- Sharp C, Pye D, Blackmore J, Eames A, Easton C, Filmer-Sankey C, Whitby K, Wilson R (2006b) *National Evaluation of Creative Partnerships: case study report: revised final version*. Slough: National Foundation for Educational Research.
- von Rossberg-Gempton IE (1998) 'Creative dance': potentiality for enhancing psychomotor, cognitive, and social-affective functioning in seniors and young children. *Dissertation Abstracts International Section A: Humanities and Social Sciences* M1-59.
- Wright R, Lindsay J, Sheel J, David R, Offord, WR (2007) Lessons learned from the National Arts and Youth Demonstration Project (longitudinal study of a Canadian after-school program). *Journal of Child and Family Studies* 14(1): 48-58.

Other references used in the text

- Angus J (2002) *A review of evaluation in community-based art for health activity in the UK*. London: NHS Health Development Agency.
- Arts Council England (2004) *The impact of the arts: some research evidence*. London: Arts Council England.
- Belfiore E (2002) Art as a means of alleviating social exclusion: does it really work? A critique of instrumental cultural policies and social impact studies in the UK. *International Journal of Cultural Policy* 8(1): 91-106.
- Belfiore E, Bennett O (2006) *Rethinking the social impact of the arts: a critical historical review*. CCPS Research Paper 9. Warwick: Centre for Cultural Policy Studies, University of Warwick.
- Belfiore E, Bennett O (2007) Determinants of impact: towards a better understanding of encounters with the arts. *Cultural Trends* 16(3): 225-275.
- Coalter F (2001) *Realising the potential of cultural services: the case for the arts*. London: Local Government Association.
- Coe R (2002) It's the effect size, stupid: what effect size is and why it is important. Paper presented at: *British Educational Research Association annual conference*, Exeter, 12-14 September.
- Daly S (2005) *Social capital and the cultural sector: literature review prepared for the Department of Culture, Media and Sport*. London: Centre for Civil Society, London School of Economics.
- Daykin N, Byrne E, Soteriou T, O'Connor S (2008) The impact of art, design and environment in mental healthcare: a systematic review of the literature. *Journal of the Royal Society for the Promotion of Health* 128(2): 85-94.

- Deasy RJ (2002) *Critical links: learning in the arts and student academic and social development*. Washington, DC: Arts Education Partnership.
- Evans G, Shaw P (2004) *The contribution of culture to regeneration in the UK: a review of evidence*. London: London Metropolitan University.
- Farrington D, Gottfredson D, Sherman L, Welsh B (2002) The Maryland Scientific Methods Scale. In: Farrington D, MacKenzie D, Sherman L, Welsh L (eds.) *Evidence based crime prevention*. London: Routledge, pages 13-21.
- Freshminds (2007) *Culture on demand: ways to engage a broader audience*. London: Department for Culture, Media and Sport.
- Galloway S (2008) *The evidence base for arts and culture policy: a brief review of selected recent literature*. Glasgow: Centre for Cultural Policy Research, University of Glasgow and Scottish Arts Council.
- Galloway S (2009) Theory-based evaluation and the social impact of the arts. *Cultural Trends* 18(2): 125–148.
- Galloway S, Birkin N, Petticrew M, Hamilton C, Bell D (2006) Quality of life and well-being: measuring the benefits of culture and sport: literature review and thinkpiece. Edinburgh: Scottish Executive.
- Gold C, Solli HP, Kruger V, Lie SA (2009) Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis. *Clinical Psychology Review* 29(3): 193-207.
- Gold C, Voracek M, Wigram T (2004) Effects of music therapy for children and adolescents with psychopathology: a meta-analysis. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 45(6): 1054-1063.
- Guetzkow J (2002) *How the arts impact communities: an introduction to the literature on arts impact studies*. Princeton, NJ: Princeton University.
- Hallam S (2009) The power of music: its impact on the intellectual, social and personal development of children and young people. http://www.ioe.ac.uk/Year_of_Music.pdf (accessed 18 January 2010).
- Health Development Agency (1999) *Art for health: a review of good practice in community-based arts projects and interventions which impact on health and wellbeing*. London: NHS Health Development Agency.
- Hughes J (2005) *Doing the arts justice: a review of research literature, practice and theory*. London. Department for Culture, Media and Sport.
- Jermyn H (2001) *The arts and social exclusion: a review prepared for the Arts Council of England*. London: Arts Council England.
- Jermyn H (2004) *The art of inclusion*. London: Arts Council England.
- Mason R, Gearon L, Valkanova Y (2006) *A systematic review of the contribution of art education to cultural learning in learners aged 5–16*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Matarasso F (1997) *Use or ornament? The social impact of participation in the arts*. Stroud: Comedia.
- McCarthy KF, Ondaatje EH, Zakaras L (2001) *Guide to the literature on participation in the arts*. Santa Monica, CA: RAND.
- McMaster B (2008) *Supporting excellence in the arts: from measurement to judgement and recent events*. London: Department for Culture, Media and Sport.

- Newman T, Curtis K, Stephens J (2001) *Do community-based arts projects result in social gains? A review of literature*. Barking: Barnardo's.
- Reeves M (2002) *Measuring the economic and social impact of the arts: a review*. London: Arts Council England.
- Ruiz J (2004) *A literature review of the evidence base for culture, the arts and sport policy*. Edinburgh: Scottish Executive.
- Shaw P (1999) *Arts and neighbourhood renewal: a literature review to inform the work of Policy Action Team 10*. London: DCMS.
- Standley J (2008) Does music instruction help children learn to read? Evidence of a meta-analysis. *Applications of Research in Music Education* 27(1): 17-32.
- Staricoff RL (2004) *Arts in health: a review of the medical literature*. London: Arts Council England.
- Timperley H, Wilson A, Barrar H, Fung Y (2007) *Teacher professional learning and development: best evidence synthesis iteration*. Wellington, New Zealand: Ministry of Education.
- White M, Angus J (2003) *Arts and adult mental health literature review: addressing the evidence base from participation in arts and cultural activities*. Durham: Centre for Arts and Humanities in Health and Medicine.
- Winner E, Cooper M (2000) Mute those claims: no evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education* 34(3/4): 11-75.

Arts in-depth review: Appendix A2

Study quality, outcome, and effect sizes

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|---|--|-----------------------|-----------------------|---------------------------------------|---------------------|
| Bigelow (1997) Medium | Transferable skills (communication apprehension): students (boys and girls) who participate in performing arts in schools vs students (boys and girls) that do not | 35 -2.900 14.962 | 33 -4.000 16.711 | 0.069 | -0.407 0.544 |
| | Transferable skills (communication apprehension): boys who participate in performing arts in schools vs boys that do not | 12 -2.800 13.965 | 18 -7.900 16.730 | 0.316 | -0.419 1.051 |
| | Transferable skills (communication apprehension): girls who participate in performing arts in schools vs girls that do not | 23 -3.000 15.861 | 15 0.600 17.067 | -0.216 | -0.868 0.437 |
| Bilhartz et al. (1999) Medium | Cognitive performance (Stanford Binet tests overall results): music treatment vs treatment as usual | 36 0.000 0.000 | 35 0.000 0.000 | 0.514 | 0.041 0.988 |
| Catterall (2007) Medium | Transferable skills (problem resolution skills): drama vs no treatment | 71 0.000 0.000 | 84 0.000 0.000 | 0.358 | 0.040 0.677 |
| | Transferable skills (meta-cognitive skills): drama vs no treatment | 71 0.000 0.000 | 84 0.000 0.000 | 0.358 | 0.040 0.677 |
| | Transferable skills (self-efficacy): drama vs no treatment | 71 0.000 0.000 | 84 0.000 0.000 | 0.408 | 0.089 0.727 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|---|---|------------------------|------------------------|---------------------------------------|---------------------|
| Catterall and Peppler (2007) Medium | Creativity (originality scale): visual arts classes vs no visual arts classes | 103 0.370 0.491 | 76 -0.040 0.517 | 0.813 | 0.505 1.121 |
| | Transferable skills (self efficacy scale): all visual sites arts vs all sites comparison | 103 0.320 0.835 | 76 -0.111 0.770 | 0.531 | 0.229 0.832 |
| Costa-Giomi (2004) Medium/High | Academic achievement/skills (language): piano vs treatment as usual | 67 0.000 0.000 | 50 0.000 0.000 | 0.486 | 0.115 0.858 |
| | Academic achievement/skills (mathematics): piano vs treatment as usual | 67 0.000 0.000 | 50 0.000 0.000 | 0.317 | -0.051 0.686 |
| | Cognitive performance (general cognitive abilities): piano instruction vs no piano | 67 0.000 0.000 | 50 0.000 0.000 | 0.483 | 0.111 0.854 |
| Fleming et al. (2004) Medium | Academic achievement/skills (reading) drama vs treatment as usual | 58 0.000 0.000 | 50 0.000 0.000 | 0.400 | 0.018 0.782 |
| | Academic achievement/skills (mathematics) drama vs treatment as usual | 58 0.000 0.000 | 50 0.000 0.000 | 0.830 | 0.435 1.225 |
| Freeman (2001) Medium/High | Transferable skills (self- concept scores): creative drama activities vs treatment as usual (no creative drama activities) | 48 0.000 0.000 | 43 0.000 0.000 | 0.238 | -0.175 0.651 |
| | Truancy rates/behaviour problems (frequency of problem behaviour): creative drama activities vs no creative drama activities | 48 0.000 0.000 | 43 0.000 0.000 | 0.099 | -0.313 0.511 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|--|---|-------------------------|-------------------------|---------------------------------------|---------------------|
| | Transferable skills (social skills score): creative drama activities vs treatment as usual (no creative drama activities) | 48 0.000 0.000 | 43 0.000 0.000 | 0.069 | -0.342 0.481 |
| Gromko (2005) Medium | Academic achievement/skills (phonemic awareness skills: letter-naming fluency): music instruction vs no music instruction | 43 9.210 10.380 | 60 7.830 12.610 | 0.117 | -0.275 0.509 |
| | Academic achievement/skills (phonemic awareness skills: phoneme- segmentation fluency): music instruction vs no music instruction | 43 26.120 16.050 | 60 15.720 13.830 | 0.698 | 0.294 1.101 |
| | Academic achievement/skills (phonemic awareness skills: nonsense-word fluency): music instruction vs no music instruction | 43 9.860 9.790 | 60 15.420 19.600 | -0.339 | -0.734 0.055 |
| Gromko and Poorman (1998) Medium | Cognitive performance (IQ) Music training vs no music training | 15 17.600 17.000 | 15 9.240 15.884 | 0.494 | -0.232 1.221 |
| | Cognitive performance (Scaled IQ) Music training vs no music training | 15 11.600 11.090 | 15 6.730 10.400 | 0.441 | -0.284 1.165 |
| Hui and Lau (2006) Medium | Creativity: participate in drama project vs non- aesthetic extra-curricular activities | 61 0.000 0.000 | 67 0.000 0.000 | 0.308 | -0.041 0.657 |
| | Creativity: participate in drama project vs non- aesthetic extra-curricular activities | 61 0.000 0.000 | 67 0.000 0.000 | 0.497 | 0.145 0.849 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|---|--|-----------------------|---------------------------|---------------------------------------|---------------------|
| | Transferable skills (communicative-expressive ability): participate in drama project vs non-aesthetic extra-curricular activities | 61 0.000 0.000 | 67 0.000 0.000 | 0.437 | 0.086 0.788 |
| Kendall et al. (2008) Medium/High | Academic achievement/skills (GCSE total points score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2003/2004) | 1,587 0.000 0.000 | 1,267,719 0.000 0.000 | 0.037 | -0.012 0.086 |
| | Academic achievement/skills (GCSE 'best 8' score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2003/2004) | 1,587 0.000 0.000 | 1,267,719 0.000 0.000 | 0.016 | -0.033 0.065 |
| | Academic achievement/skills (GCSE English): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2003/2004) | 1,587 0.000 0.000 | 1,267,719 0.000 0.000 | -0.011 | -0.060 0.038 |
| | Academic achievement/skills (GCSE mathematics): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2003/2004) | 1,587 0.000 0.000 | 1,267,719 0.000 0.000 | -0.009 | -0.058 0.040 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|-------------------|--|-----------------------------|---------------------------------|---------------------------------------|---------------------|
| | Academic achievement/ skills (GCSE science): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2003/2004) | 1,587 0.000 0.000 | 1,267,719 0.000 0.000 | 0.051 | 0.002 0.100 |
| | Academic achievement/ skills (average KS2 score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 8,670 0.000 0.000 | 1,005,105 0.000 0.000 | -0.001 | -0.022 0.020 |
| | Academic achievement/ skills (KS2 English score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 8,670 0.000 0.000 | 1,005,105 0.000 0.000 | 0.016 | -0.005 0.037 |
| | Academic achievement/ skills (KS2 mathematics score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 8,670 0.000 0.000 | 1,005,105 0.000 0.000 | -0.001 | -0.022 0.020 |
| | Academic achievement/ skills (KS2 science score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 8,670 0.000 0.000 | 1,005,105 0.000 0.000 | -0.019 | -0.040 0.002 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|-------------------|--|-----------------------------|---------------------------------|---------------------------------------|---------------------|
| | Academic achievement/ skills (average KS3 score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 6,493 0.000 0.000 | 1,104,907 0.000 0.000 | 0.062 | 0.038 0.086 |
| | Academic achievement/ skills (KS3 English score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 6,493 0.000 0.000 | 1,104,907 0.000 0.000 | 0.081 | 0.057 0.105 |
| | Academic achievement/ skills (KS3 mathematics score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 6,493 0.000 0.000 | 1,104,907 0.000 0.000 | 0.048 | 0.024 0.072 |
| | Academic achievement/ skills (KS3 science score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally | 6,493 0.000 0.000 | 1,104,907 0.000 0.000 | 0.054 | 0.030 0.078 |
| | Academic achievement/ skills (GCSE total point score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2005/2006) | 5,188 0.000 0.000 | 1,081,248 0.000 0.000 | 0.101 | 0.074 0.128 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|---|--|-----------------------------|---------------------------------|---------------------------------------|---------------------|
| | Academic achievement/skills (GCSE 'best 8' point score): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2005/2006) | 5,188 0.000 0.000 | 1,081,248 0.000 0.000 | 0.077 | 0.050 0.104 |
| | Academic achievement/skills (GCSE mathematics): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2005/2006) | 5,188 0.000 0.000 | 1,081,248 0.000 0.000 | 0.028 | 0.001 0.055 |
| | Academic achievement/skills (GCSE English): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2005/2006) | 5,188 0.000 0.000 | 1,081,248 0.000 0.000 | 0.069 | 0.042 0.096 |
| | Academic achievement/skills (GCSE science): overall difference between young people known to have attended 'Creative Partnerships' and similar young people nationally (KS4: 2005/2006) | 5,188 0.000 0.000 | 1,081,248 0.000 0.000 | 0.064 | 0.037 0.091 |
| Kim et al. (2008) Medium/High | Transferable skills (joint attention skills and pro-social behaviours): music therapy vs play sessions (time 3 minus time 1) | 8 0.000 0.000 | 7 0.000 0.000 | 0.744 | -0.317 1.804 |
| | Transferable skills (joint attention skills and pro-social behaviours): music therapy vs play sessions (time 2 minus time 1) | 8 0.000 0.000 | 7 0.000 0.000 | 0.913 | -0.170 1.996 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|--|---|-----------------------|-----------------------|---------------------------------------|---------------------|
| Krahe and Knappert (2009) Medium | Transferable skills (self-protective skills): attend live drama vs no treatment (time 2 minus time 1) | 44 6.710 6.480 | 49 1.580 7.030 | 0.751 | 0.329 1.173 |
| | Transferable skills (self-protective skills): attend live drama vs no treatment (time 3 minus time 1) | 44 6.000 7.290 | 49 0.020 9.250 | 0.708 | 0.287 1.128 |
| McMahon et al. (2003) Medium | Academic achievement/skills (consonant sound recognition) Basic Reading through Dance (BRD) vs no BRD | 290 0.000 0.000 | 338 0.000 0.000 | 0.589 | 0.429 0.750 |
| | Academic achievement/skills (overall phoneme) Basic Reading through Dance (BRD) vs no BRD | 242 0.000 0.000 | 274 0.000 0.000 | 0.619 | 0.442 0.796 |
| | Academic achievement/skills (vowel recognition) Basic Reading through Dance (BRD) vs no BRD | 266 0.000 0.000 | 338 0.000 0.000 | 0.360 | 0.198 0.521 |
| Merrell (2005) Medium | Transferable skills (actual reporting of bullying): drama-based social skills programme vs treatment as usual (Freshman Seminars) | 30 0.240 0.570 | 30 0.280 1.746 | -0.030 | -0.536 0.476 |
| | Transferable skills (direct intervention in bullying situations): drama-based social skills programme vs treatment as usual (Freshman Seminars) | 30 0.000 0.000 | 30 0.000 0.000 | 0.158 | -0.349 0.665 |
| | Transferable skills (attitudes to bullying): drama-based social skills programme vs treatment as usual (Freshman Seminars) | 30 0.000 0.000 | 30 0.000 0.000 | 0.355 | -0.155 0.866 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|--|--|-------------------------|------------------------|---------------------------------------|---------------------|
| Orsmond and Miller (1999) Medium | Cognitive performance (Peabody Picture Vocabulary Test): music training vs no music training | 29 5.860 17.911 | 29 6.380 26.348 | -0.023 | -0.538 0.492 |
| | Cognitive performance (Developmental Test of visual-motor Integration): music training vs no music training | 29 1.780 2.984 | 29 0.270 2.557 | 0.536 | 0.012 1.060 |
| | Cognitive performance (Puzzle solving subtest): music training vs no music training | 29 10.050 16.737 | 29 5.710 20.909 | 0.226 | -0.290 0.742 |
| | Cognitive performance (Preschool Embedded Figures Test): music training vs no music training | 29 -0.230 2.325 | 29 -0.530 2.538 | 0.122 | -0.394 0.637 |
| Piro and Ortiz (2009) Medium | Academic achievement/ skills (early reading skills/vocabulary): keyboard lessons vs no keyboard lessons | 46 | 57 | 1.28 | 0.85 1.7 |
| | Academic achievement/ skills (early reading skills/verbal sequencing): keyboard lessons vs no keyboard lessons | 46 | 57 | 1.5 | 1.1 2.0 |
| Rauscher and Zupan (2000) Medium | Cognitive performance, music instruction featuring the keyboard vs treatment as usual (no keyboard); Puzzle solving task | 34 7.450 5.210 | 28 2.940 3.170 | 1.010 | 0.479 1.541 |
| | Cognitive performance: music instruction featuring the keyboard vs treatment as usual (no keyboard access); Pictorial Memory Test baseline to 8 months | 34 1.500 1.172 | 28 0.570 1.136 | 0.794 | 0.275 1.314 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|---|--|------------------------------|--------------------------|---------------------------------------|---------------------|
| | Cognitive performance; music instruction featuring the keyboard vs treatment as usual (no keyboard access); Block Building test pretest-8 months | 28 - 18.960 45.100 | 34 -49.960 42.550 | 0.700 | 0.185 1.215 |
| | Cognitive performance; music instruction featuring the keyboard vs treatment as usual (no keyboard access); Puzzle solving pretest-4 months | 34 4.650 4.341 | 28 1.820 2.893 | 0.743 | 0.226 1.260 |
| | Cognitive performance; music instruction featuring the keyboard vs treatment as usual (no keyboard access); Block Building pretest-4 months | 28 -3.120 46.599 | 34 -37.940 44.513 | 0.756 | 0.239 1.274 |
| | Cognitive performance; music instruction featuring the keyboard vs treatment as usual (no keyboard access); Pictorial Memory pretest-4 months | 34 0.940 0.995 | 28 -0.290 1.282 | 1.072 | 0.537 1.607 |
| Rauscher et al. (1997) Medium | Cognitive performance: Piano vs as usual (OA scores immediate post test) | 34 3.620 2.770 | 14 0.500 2.550 | 1.133 | 0.470 1.795 |
| | Cognitive performance: structured singing lesson vs activity as usual | 10 0.300 3.560 | 14 0.500 2.550 | -0.064 | -0.876 0.747 |
| Register (2004) Medium | Academic achievement/ skills (emergent literacy): changes in scores for letter naming: music vs control | 22 8.050 17.526 | 20 4.050 15.090 | 0.239 | -0.369 0.847 |
| | Academic achievement/ skills (emergent literacy): changes in scores for initial sounds fluency: music vs control | 22 8.270 7.712 | 20 6.250 5.981 | 0.285 | -0.323 0.894 |

| Author Quality | Outcome ¹ | Group 1 N M SD | Group 2 N M SD | Hedges' g (corrected) ² | CI lower upper |
|--|--|------------------------|------------------------|--|---------------------|
| | Academic achievement/ skills (emergent literacy): changes in scores for Test of Early Reading Ability 3rd Edition: music vs control | 22 4.130 7.355 | 20 5.600 9.118 | -0.175 | -0.782 0.432 |
| Schellenberg (2004) Medium/High | Cognitive performance (IQ): keyboard lessons vs no lessons | 36 6.100 11.121 | 36 3.900 9.801 | 0.208 | -0.256 0.671 |
| | Cognitive performance (IQ): voice lessons vs no lessons | 36 7.600 11.842 | 36 3.900 9.802 | 0.337 | -0.129 0.802 |
| | Cognitive performance (IQ): drama lessons vs no lessons | 36 5.100 13.701 | 36 3.900 9.802 | 0.100 | -0.363 0.562 |
| Standley and Hughes (1997) Medium | Academic achievement/ skills: developmental writing and language skills score: music lessons with emphasis on pre-reading and writing vs no music lessons | 17 0.000 0.000 | 15 0.000 0.000 | 0.907 | 0.173 1.640 |
| von Rossberg- Gempton et al. (1999) Medium | Cognitive abilities: creative dance vs physical education | 20 14.421 6.021 | 13 10.923 5.346 | 0.592 | -0.123 1.306 |
| | Transferable skills: positive affect/happiness: dance vs physical education (12 weeks) | 20 -0.308 0.855 | 13 -0.125 0.641 | -0.229 (for this instrument a negative effect size = a favourable outcome) | -0.930 0.472 |
| | Transferable skills: affective skills (self- concept): dance vs physical education (pre- post) | 20 -0.526 4.221 | 13 0.000 1.528 | 0.15 | 0.000 0.000 |

1. Where a zero is shown in the mean column, it means that the effect size was calculated from another value given by the author in the paper, e.g., a t-value.

2. Where a zero is shown in the column, it means that the effect size is smaller than 0.000

Arts in-depth review: Appendix A3

Arts studies structured abstracts

| Item | Study | Sample | Activity/Intervention |
|---|--|---|--|
| Bigelow (1997) Overall weight of evidence: Medium | <p>What are the broad aims of the study? 'The purpose of this study is to examine the relationship between (a) participation in the performing arts within school curricula and (b) levels of communication apprehension.' (p.4)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills: <i>1. Communication skills</i></p> <p>Which methods were used to collect the data? Personal Report of Communication Apprehension-24 (PRCA-24)</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants Alpha school: 35 Omega school: 33</p> <p>Age of participants 11-15 years</p> <p>Type of educational institution attended Secondary school</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Arts (multi-component): performing arts course (instrumental and vocal music, drama and dance) (p.11)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention One semester (p.48)</p> <p>Intensity of the activity/intervention 1-2 per week (assumed)</p> <p>What treatment/intervention did the control/comparison group receive? No treatment: core elective classes (but not performing arts classes)</p> |
| Bilhartz et al. (1999) | <p>What are the broad aims of the study? To examine the relationship between</p> | <p>In which country/countries was the study conducted?</p> | <p>Which sectors does the engagement relate to?</p> |

| Item | Study | Sample | Activity/Intervention |
|------------------------------------|---|---|--|
| Overall weight of evidence: Medium | <p>participation in a structured music curriculum and cognitive development.</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Cognitive performance</p> <p>Which methods were used to collect the data? 1. Parents or guardians were asked to complete a questionnaire about their children. (p.618) 2. Young Child Music Skills Assessment (MSA) <i>(not relevant to this review)</i> 3. Stanford-Binet Intelligence Scale, fourth edition (SB)</p> | <p>USA</p> <p>Number of participants 71 participants: 36 (experimental treatment group) 35 (control group) (p.618)</p> <p>Age of participants 0-5 years</p> <p>Type of educational institution attended Pre-school/nursery/kindergarten</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Arts</p> <p>Does the activity/intervention have a formal name? Yes: Kindermusik for the Young Child Year 1 Pilot Programme</p> <p>Arts participation Arts (multi-component): 'The weekly lessons for the age group in this study involve vocal exploration and matching pitch, singing, playing percussion instruments and the glockenspiel, exploring and notating basic rhythms, learning to read and write music on a treble staff, composing, and developing coordination and balance through movement.' (p.620)</p> <p>Duration of the activity/intervention 30 weeks (p.620)</p> <p>Intensity of the activity/intervention 75 minutes once weekly (p.620)</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: '..children in the control group attended their respective preschools but received no additional in-class music treatment.' (p.620)</p> |
| Catterall | What are the broad aims of the study? | In which country/countries was the | Which sectors does the engagement |

| Item | Study | Sample | Activity/Intervention |
|--|---|--|--|
| (2007) Overall weight of evidence: Medium | <p>'This article explores learning in drama through contemporary theories of knowledge acquisition' (p.2) 'A primary goal was enhancing prosocial behavior....' (p.9)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills <i>1. Conditions and processes of learning (meta-cognitive skills, self-efficacy, general outlook)</i> <i>2. Pro-social changes in behaviour (ability to work with others when disagreeing, ability to work effectively in groups, problem resolution skills) (pp. 17-18)</i></p> <p>Which methods were used to collect the data? 'The principal means of generating and gathering data was through surveys administered to all students prior to the program and again after completion of the program...The items in the survey instrument were supplemented with questions based on standard instruments designed to assess student attitudes and motivation.' (p.10)</p> | <p>study conducted? USA</p> <p>Number of participants 71 (intervention), 84 (control) (p.10)</p> <p>Age of participants 11-15 years</p> <p>Type of educational institution attended Secondary school</p> <p>Sex of participants Not stated</p> <p>Are outcomes reported for any of the priority groups? BME, low SES (high proportion of participants from these groups)</p> | <p>relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: the 'School Project'.</p> <p>Arts attendance and participation Arts (multi-component): 'The program used theatre, movement, and writing, as well as voice, drawing, and visual arts exercises as building blocks for students to write and perform original plays. On one of the appointed days, all students were taken to see a professional play' (p.8)</p> <p>What is/are the setting(s) of the activity/intervention? School-based extra-curricular clubs</p> <p>Duration of the activity/intervention 24 weeks</p> <p>Intensity of the activity/intervention 90 minute workshops once a week</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual</p> |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|---|
| Catterall and Peppler (2007) Overall weight of evidence: Medium | <p>What are the broad aims of the study? The broad aims were to investigate 'the effects of rich, sustained visual arts instruction on inner city 9-year-olds in two major US cities.' (p.543)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills: <ol style="list-style-type: none"> 1. General self-concept 2. Self-efficacy 3. Internal attributions for success 4. Perceived number of future choices Creativity: <ol style="list-style-type: none"> 1. Originality 2. Elaboration 3. Flexibility 4. Fluency </p> <p>Which methods were used to collect the data? 'Survey items were drawn to establish multi-item scales</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 179 students took part (103 who attended ICA or COCA classes and 76 comparison class students). (p.552)</p> <p>Age of participants 6-10 years</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants There was one treatment (arts) and one control, but two different sites were used in the study (Los Angeles, St. Louis).</p> <p>Are outcomes reported for any of the priority groups? BME, low SES (participants almost exclusively from these groups)</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: 1. Inner City Arts (ICA) based in Los Angeles, 2. Centre of Contemporary Arts (COCA) based in St. Louis</p> <p>Arts participation Arts (multi-component): At ICA the activities were drawing, painting and sculpting. The programme at COCA involved creating visual art and writing poetry.</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours) Other educational setting</p> <p>Duration of the activity/intervention 20 weeks (ICA) 30 weeks (COCA)</p> <p>Intensity of the activity/intervention About 90 minutes, twice per week (ICA) once per week for one-hour (COCA)</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual (assumed they</p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| | | | received standard art classes) |
| Costa-Giomi (2004) Linked report: Costa-Giomi (1999) Overall weight of evidence: Medium/High | <p>What are the broad aims of the study? 'to investigate the effects of piano instruction on children's development'</p> <p>What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>If academic achievement is measured, what subjects does the study focus on? Literacy, Numeracy</p> <p>Which methods were used to collect the data? 1. Language and mathematics subtests: Level 14 of the Canadian Achievement Test 2 (CAT2) 2. Coopersmith Self-Esteem Inventories (long form) (<i>not relevant to this review</i>) 3. Level E of the Developing Cognitive Abilities Test (DCAT) 4. Report cards for English, French, music and mathematics 5. Fine motor subtests of the Bruininks-Oseretsky Test of Motor Proficiency (<i>not</i></p> | <p>In which country/countries was the study conducted? Canada</p> <p>Number of participants 117 children (58 girls, 59 boys) (p.142)</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants 'The 117 children (58 girls and 59 boys) selected to participate in the study had never participated in formal music instruction, did not have a piano at home, and their family income was below \$40,000 Canadian (\$30,000 US at the time of the investigation) per annum. Approximately 25 percent of the children had unemployed parents and 30 percent lived with a single parent.' (p.142)</p> <p>Are outcomes reported for any of the priority groups?</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Play a musical instrument (<i>piano/keyboard instruction</i>)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 'Each child in the experimental group received.....three years of piano instruction and an acoustic piano.' p143</p> <p>Intensity of the activity/intervention 1-2 per week: 'The lessons were 30 minutes long during the first two years and 45 minutes during the third year.' p143</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual</p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|---|---|
| | <i>relevant to this review)</i> | Low SES (intervention targeted at this group) | |
| <p>Fleming et al. (2004)</p> <p>Linked report: Merrell and Tymms (2002)</p> <p>Overall weight of evidence: Medium</p> | <p>What are the broad aims of the study? 'This article reports on research which examined the impact of The National Theatre's Transformation drama project on young pupils' reading, mathematics, attitude, self-concept and creative writing in primary schools.' (p.177)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills Attitude to learning <i>Attitude to mathematics, reading and school (not used in outcome measures as data not reported to calculate effect sizes)</i></p> | <p>In which country/countries was the study conducted? UK</p> <p>Number of participants 165 students</p> <p>Age of participants 6-10 years</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants The primary schools were in the Shadwell and Limehouse areas of the East End of London (Borough of Tower Hamlets). A large number of the pupils were learning English as an additional language. (p.181). There were 2 intervention schools and 2 control schools.</p> <p>Are outcomes reported for any of the priority groups? Unclear/BME (many children were learning English as an additional language)</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: The National Theatre's Transformation drama project.</p> <p>Arts attendance and participation Arts (multi-component): 1. Write plays, perform plays 2. Part of the project involved three visits to the theatre to watch professional performances. (p.182)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours) Arts setting (e.g. theatre/gallery)</p> <p>Duration of the activity/intervention Just under 2 years. 'All Year 3 pupils were assessed in September 1999 ... They were then re-assessed at the end of Year 4 after two years of Transformation interventions.' (p.183)</p> <p>Intensity of the activity/intervention "The broad pattern for working with the children was similar each year. It took</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|--|
| | | | <p>place over two terms with serial workshops in the first term and a two-week block in the second. During the second stage, five drama workshops were conducted leading up to a celebratory sharing of the work in a venue outside schools. Part of the project involved three visits to the theatre to watch professional performances." (pp.181-2)</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: standard drama lessons (assumed)</p> |
| <p>Freeman (2001)</p> <p>Overall weight of evidence: Medium/High</p> | <p>What are the broad aims of the study? 'The purpose of this study was to determine the effects of creative drama activities on the self-concept, behaviour, and social skills of third and fourth year students.' (p.2)</p> <p>What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5)</p> <p>What learning outcomes does the study measure/report? Transferable skills 1. <i>Social skills</i> 2. <i>Self-concept</i> Truancy rates/behaviour problems</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 237 subjects participated: 119 from grade 3, 118 from grade 4; 120 in the intervention group, 117 controls</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Arts (multi-component): Creative drama activities</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 18 weeks</p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| | <i>Behaviour</i> Which methods were used to collect the data? 1. The Student Self-Concept Scale (SSCS) was used to measure student self-concept. 2. The Social Skills Rating System (SSRS) was used to measure social skills and problem behaviour. 3. Behaviour was also measured using computer records of office referrals from the study year and the prior year. | Mixed ethnicity and SES Are outcomes reported for any of the priority groups? No | Intensity of the activity/intervention One day per week for 40 minutes What treatment/intervention did the control/comparison group receive? Treatment as usual: 'regularly scheduled general music classes' (p.iv) |
| Gromko (2005) Overall weight of evidence: Medium | What are the broad aims of the study? 'The purpose of this study was to determine whether music instruction was related to significant gains in the development of young children's phonemic awareness, particularly in their phoneme-segmentation fluency.' (abstract) What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3) What learning outcomes does the study measure/report? Academic achievement/skills <i>Phonemic awareness skills</i> Which methods were used to collect | In which country/countries was the study conducted? USA Number of participants Details <i>43 intervention, 60 control (103 total)</i> Age of participants 0-5 years Type of educational institution attended kindergarten at an elementary school Sex of participants Mixed sex Are outcomes reported for any of the priority groups? | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? No Arts participation Arts (multi-component): Music instruction involving singing, body percussion/kinaesthetic movement, playing instruments What is/are the setting(s) of the activity/intervention? School (in school hours) Duration of the activity/intervention |

| Item | Study | Sample | Activity/Intervention |
|---|--|---|--|
| | the data? <i>LNF, PSF, and NMF subtests of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) test</i> | No | <i>4 months</i> Intensity of the activity/intervention <i>Once a week for 30 minutes</i> What treatment/intervention did the control/comparison group receive? No treatment |
| Gromko and Poorman (1998) Overall weight of evidence: Medium | What are the broad aims of the study? The purpose of this study is to investigate the effect of music training on preschoolers' performance IQ using 5 spatial-temporal tasks. What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3) What learning outcomes does the study measure/report? Cognitive performance Which methods were used to collect the data? WPPSI-R (Wechsler Preschool and primary intelligence scale- revised) | In which country/countries was the study conducted? USA Number of participants 34 (17 treatment, 17 control) at recruitment 4 dropped out so 30 in study Age of participants 0-5 years Type of educational institution attended Pre-school/nursery/kindergarten Sex of participants Not stated Any other important features of the participants Our subjects were drawn from a private Montessori school where all children learn with an intellectually stimulating environment, receive a 'traditional' music | Which sectors does the engagement relate to? Arts Arts participation Arts (multi-component): Children a) sang the new song several times, b) accompanied the singing with body percussion that 'painted' the melody in the air, c) took turns playing the simplified version of the song on songbells or handchimes, d) made a picture of the song using round stickers on paper, e) followed a tactile touch chart that outlined the contour of the song.' Two familiar songs were danced and sung. Every family purchased a 20 note set of songbells ... and kept at home for practice. (p.176) What is/are the setting(s) of the activity/intervention? School (in school hours) |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|--|
| | | <p>class. ...and benefit from outside activities provided by their parents.</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Duration of the activity/intervention 24 Tuesdays</p> <p>Intensity of the activity/intervention 1/2 hour, once a week</p> <p>What treatment/intervention did the control/comparison group receive? Not stated</p> |
| <p>Hui and Lau (2006)</p> <p>Overall weight of evidence: Medium</p> | <p>What are the broad aims of the study? To examine 'the effect of drama education on the psychological development of grades 1 and 4 students.' (abstract)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills: 1. <i>Communicative-expressive ability</i> Creativity: 1. <i>Thinking and drawing</i></p> <p>Which methods were used to collect the data? 1. Form A of the Wallach-Kogan creativity tests (WKCT) 2. Tests for creative thinking-drawing production (TCT-DP) 3. A story-telling test (STT) was designed</p> | <p>In which country/countries was the study conducted? Asia</p> <p>Number of participants It is reported that a total of 126 children were assigned to the experimental group and the control group included 69 children (p.35). However, Table 1 on p.37 suggests that only 61 (intervention) and 67 (control) actually took part.</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Arts (multi-component): drama project involving puppet making and creative drama (improvisation and story creation) (p.36)</p> <p>What is/are the setting(s) of the activity/intervention? School-based extra-curricular clubs</p> <p>Duration of the activity/intervention 16 weeks</p> <p>Intensity of the activity/intervention One day each week</p> <p>What treatment/intervention did the</p> |

| Item | Study | Sample | Activity/Intervention |
|--|---|---|--|
| | and used to measure students' communicative-expressive ability. 4. Observation/video-taping | priority groups? No | control/comparison group receive? Alternative intervention: 'non-aesthetic extra-curricular activities, such as ball playing' (p.36) |
| Kendall et al. (2008) Linked reports: Eames et al. (2006); Kendall et al. (2008); Sharp et al. (2006a); Sharp et al. (2006b) Overall weight of evidence: Medium/High | What are the broad aims of the study? 'The main aim of this report is to explore the relationship between taking part in Creative Partnerships activities, or attending a school which took part in the Creative Partnerships initiative, and academic attainment.' (p.8) What was the design of the evaluation? Well-matched comparison group study (Maryland Scale 4) What learning outcomes does the study measure/report? Academic achievement/skills <i>Key Stage 2 and 3 assessments</i> <i>GCSEs</i> If academic achievement is measured, what subjects does the study focus on? Literacy Numeracy Science Which methods were used to collect the data? | In which country/countries was the study conducted? UK Number of participants Number of young people in the analysis (Table A1, p.29) Key stage 2: 8,670* 12,102** 1,005,105*** Key stage 3: 6,493* 24,883** 1,104,907*** Key stage 4: 5,188* 23,921** 1,081,248*** * Young people known to have attended Creative Partnership activities **All young people in schools involved with Creative Partnerships in Phase 1 ***All young people nationally Number of schools in the analysis (Table A2, p.30) Key stage 2: 158* 14,126** Key stage 3: 73* 3,053** Key stage 4: 73* 3,034** *Schools involved with Phase 1 of Creative Partnerships **All schools nationally Age of participants | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? Yes: Creative Partnerships Arts attendance and participation Arts (multi-component): The programme supports innovative, long-term partnerships between schools and creative professionals including artists, performers, architects, multimedia developers and scientists. What is/are the setting(s) of the activity/intervention? School (in school hours) Duration of the activity/intervention Not stated Intensity of the activity/intervention Not stated What treatment/intervention did the control/comparison group receive? |

| Item | Study | Sample | Activity/Intervention |
|--|---|--|--|
| | 1. National Pupil Database (NPD) data 2. Attendance data collected during the national evaluation | 6-10 years 11-15 years 16-18 years Type of educational institution attended Primary school Secondary school Sex of participants Mixed sex Are outcomes reported for any of the priority groups? Unclear/low SES (intervention was targeted at 'disadvantaged communities') | No treatment: The models effectively compared the progress of three mutually exclusive groups of young people: those known to have attended Creative Partnerships activities those in Creative Partnerships schools, but not known to have attended Creative Partnerships activities those in non-Creative Partnerships schools. |
| Kim et al. (2008) Overall weight of evidence: Medium/High | What are the broad aims of the study? 'to investigate the effects of improvisational music therapy on joint attention behaviours in pre-school children with autism.' (Abstract) What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5) What learning outcomes does the study measure/report? 1. Joint attention skills 2. Pro-social behaviours Which methods were used to collect the data? | In which country/countries was the study conducted? Asia Number of participants 15 at recruitment 10 in final study Age of participants 0-5 years Type of educational institution attended Not stated Sex of participants Male | Which sectors does the engagement relate to? Arts Arts participation Arts (multi-component): improvisational music therapy Duration of the activity/intervention 'Due to holidays and sick leave, it took the participants between 7 and 8 months to complete the 24 session program.' (pp.1759-60) Intensity of the activity/intervention Weekly, 30 minute sessions |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|--|
| | 1. PDDBI (Pervasive Developmental Disorder Behaviour Inventory) 2. ESCS (a structured toy play assessment measuring non-verbal social communication skills) 3. Video | Any other important features of the participants Participants had autism. Those that remained after drop-out were all boys. Are outcomes reported for any of the priority groups? Limiting disability (children with autism) | What treatment/intervention did the control/comparison group receive? Treatment as usual: play session with toys |
| Krahe and Knappert (2009) Overall weight of evidence: Medium | What are the broad aims of the study? To examine the efficacy of an intervention targeting first and second graders ... the first test of the efficacy of a theatre play called (No) Child's Play in promoting children's self-protective skills in terms of the understanding of situations potentially leading to abuse and in recognising appropriate behavioural responses in interactions with adults. (p.322) What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5) What learning outcomes does the study measure/report? Transferable skills: 1. <i>Self-protective skills</i> Which methods were used to collect the data? Eight short scenarios were developed to measure gains in self-protective skills. | In which country/countries was the study conducted? Europe Number of participants 'A total of 148 first and second graders (78 girls; average age 7.55 years ... range 6.1–9.1 years) participated in the study ... In total, there were 44 participants in the LIVE group, 55 participants in the DVD group and 49 in the control group.' (p.323) Age of participants 6-10 years Type of educational institution attended Primary school Sex of participants Mixed sex Are outcomes reported for any of the priority groups? | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? Yes: (No) Child's Play Arts attendance Theatre-based intervention (i.e. live performance of a play) to prevent sexual abuse What is/are the setting(s) of the activity/intervention? Arts setting: theatre Duration of the activity/intervention One day or less Intensity of the activity/intervention Once What treatment/intervention did the control/comparison group receive? |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|---|
| | The scenarios were followed by questions (answered by students) with four-point rating scale. (pp.324-5) | No | No treatment |
| McMahon et al. (2003) Overall weight of evidence: Medium | <p>What are the broad aims of the study? 'This study evaluates the effectiveness of an arts-based educational program, Basic Reading Through Dance.' (Abstract)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>If academic achievement is measured, what subjects does the study focus on? Literacy</p> <p>Which methods were used to collect the data? 1. Read America's PhonoGraphix Test 2. Code Knowledge 3. Phoneme Segmentation</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 721 first-grade students participated in the study. 'At post-test, 630 students were assessed, with 293 in the experimental group and 337 in the control group.' (p.111)</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? BME, low SES (participants came from schools that served predominantly African American populations from poverty-level families)</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: Basic Reading Through Dance programme</p> <p>Arts participation Dance programme 'targeted specifically at improving basic reading skills in the areas of decoding and phoneme-grapheme relationships.' (p.110)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 10 weeks</p> <p>Intensity of the activity/intervention 40 minute sessions, twice a week (students were exposed to a total of 13.3 hours of dance-based reading instruction, p.112)</p> <p>What treatment/intervention did the control/comparison group receive?</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|---|
| | | | Treatment as usual: usual reading instruction |
| Merrell (2005) Overall weight of evidence: Medium | <p>What are the broad aims of the study? The purpose of this study is to test the hypothesis that drama related to bullying improves the bystander's response to bullying situations</p> <p>What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5)</p> <p>What learning outcomes does the study measure/report? Transferable skills 1. <i>Actual reporting of bullying</i> 2. <i>Direct intervention in bullying situations</i> 3. <i>Attitudes about bullying</i></p> <p>Which methods were used to collect the data? 1. Pre-Intervention Bullying Questionnaire Format A 2. Post-Intervention Bullying Questionnaire Format B 3. Reflection cards</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants Experimental (n=30) and control (n=30) groups. After the study began, one subject in experimental group was temporarily suspended, and three in the control group decided not to participate. Therefore, a total of 56 ninth graders (29 in intervention, 27 in control) were included. (p.81)</p> <p>Age of participants 11-15 years</p> <p>Type of educational institution attended Secondary school</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? BME (large proportion of sample)</p> | <p>Which sectors does the engagement relate to? Art</p> <p>Does the activity/intervention have a formal name? Yes: The 5 W's of Bullying Intervention</p> <p>Arts participation Arts (multi-component): Drama-based social-skills programme</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention Unclear</p> <p>Intensity of the activity/intervention 7 sessions at 45 minutes</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: 'The control group completed seven sessions of the regular Freshman Seminars, a programme focused on high school credits, careers and study skills given to all nine graders in the school.'</p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| | | | (abstract) |
| Orsmond and Miller (1999) Overall weight of evidence: Medium | <p>What are the broad aims of the study? 'This study investigated the effects of early music instruction on music perception skills as well as specific areas of cognitive development (non-verbal/spatial abilities).' (abstract)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Cognitive performance: <i>A variety of cognitive measures (receptive vocabulary, visual-motor integration, an embedded figures task, and puzzles)</i></p> <p>Which methods were used to collect the data? Questionnaire given to parents which collected background information</p> <p>Tests to measure cognitive skills (see pp.22-23 for further details): 1. Peabody Picture Vocabulary Test - Revised (PPVT) 2. Development Test of visual-motor integration, Third Revision (VMI) 3. 20 picture puzzles modelled after the Puzzle-Solving subtest of the McCarthy</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants <i>29 intervention, 29 control (58 total)</i></p> <p>Age of participants 0-5 years 6-10 years <i>Range: 3 years 8 months to 6 years 8 months</i></p> <p>Type of educational institution attended Pre-school/nursery/kindergarten</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: Suzuki music instruction</p> <p>Arts participation Suzuki music instruction: i.e., instrument lessons (types of instrument not detailed)</p> <p>What is/are the setting(s) of the activity/intervention? Arts setting: (private) music centres (no further details)</p> <p>Duration of the activity/intervention Children were tested after 4 months of music lessons.</p> <p>Intensity of the activity/intervention Not stated</p> <p>What treatment/intervention did the control/comparison group receive? No treatment</p> |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|--|
| | Scales of Children's Abilities 4. Preschool Embedded Figures Test (PEFT) | | |
| Piro and Ortiz (2009) Overall weight of evidence: Medium | <p>What are the broad aims of the study? 'The major aim of this quasi-experimental study was to examine the effects of a scaffolded music instruction program on the vocabulary and verbal sequencing skills of two cohorts of second-grade students.' (abstract)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills <i>Vocabulary and verbal sequencing skills</i></p> <p>Which methods were used to collect the data? Students were administered two subtests of the Meeker Structure of Intellect (SOI) – Form L. The first subtest used was Vocabulary. The second subtest was Verbal Sequencing.</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 46 intervention, 57 control</p> <p>Age of participants 6-10 years: <i>2nd grade</i></p> <p>Type of educational institution attended 'Participants in both treatment and control groups attended two large public elementary schools found in the same middle-class area of New York City.' (p.5)</p> <p>Sex of participants Mixed sex</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Music training (keyboard lessons) as a part of the general school curriculum. The 'scaffolded' music instruction was delivered parallel with the school's balanced literacy programme, which included daily lessons in reading, writing, speaking, and listening.</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention All participants were individually pre- and post-tested at the start and close of a standard 10-month school year. Participants had taken part in activities for 3 consecutive years.</p> <p>Intensity of the activity/intervention</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|---|---|
| | | | Twice a week (for periods of 40-45 minutes) What treatment/intervention did the control/comparison group receive? No treatment |
| Rauscher et al. (1997) Overall weight of evidence: Medium | What are the broad aims of the study? To examine whether musical learning improves long-term spatial temporal reasoning ability in early years children. (pp.2-3) What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3) What learning outcomes does the study measure/report? Cognitive performance Which methods were used to collect the data? Object assembly task of the Wechsler preschool and primary scale of intelligence-revised (WPPPSI-R) | In which country/countries was the study conducted? USA Number of participants 111 initially recruited, 33 children withdrew from pre-schools during the course of the study leaving 78 34 keyboard group 10 singing group 20 computer group no lessons = 14 Age of participants 0-5 years Type of educational institution attended Pre-school/nursery/kindergarten Sex of participants Mixed sex Are outcomes reported for any of the priority groups? No | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? No Arts participation 1. Learning/playing a musical instrument (<i>keyboard instruction</i>) 2. Singing lessons What is/are the setting(s) of the activity/intervention? School (in school hours) Duration of the activity/intervention 6 months (SA school) 8 months (LB and WC schools) Intensity of the activity/intervention Keyboard: once a week (LB and WC schools) twice a week (SA school) Singing: 5 days a week (p.3) |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|---|
| | | | What treatment/intervention did the control/comparison group receive? Treatment as usual: no lessons |
| Rauscher and Zupan (2000) Overall weight of evidence: Medium | What are the broad aims of the study? The purpose of this study was to determine the effects of classroom music instruction featuring the keyboard on the spatial-temporal reasoning of kindergarten children What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3) What learning outcomes does the study measure/report? Cognitive performance Which methods were used to collect the data? Spatial-temporal tasks involving (a) puzzle solving; (b) block building; (c) pictorial memory | In which country/countries was the study conducted? USA Number of participants 62 keyboard n=34 no music n=28 Age of participants 0-5 years 6-10 years Type of educational institution attended Pre-school/nursery/kindergarten Sex of participants Mixed sex Any other important features of the participants Mixed ethnicity Are outcomes reported for any of the priority groups? No | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? No Arts participation Play a musical instrument (<i>keyboard instruction</i>) What is/are the setting(s) of the activity/intervention? School (in school hours) Duration of the activity/intervention Unclear Intensity of the activity/intervention Twice a week, 20 minutes What treatment/intervention did the control/comparison group receive? No treatment |
| Register (2004) | What are the broad aims of the study? 'to examine the effects of a music therapy program designed to teach reading skills | In which country/countries was the study conducted? | Which sectors does the engagement relate to? |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|--|
| <p>Linked report: Register (2003)</p> <p>Overall weight of evidence: Medium</p> | <p>on the early literacy behaviours of Kindergarten children from a low socioeconomic background' (p. 7)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>If academic achievement is measured, what subjects does the study focus on? Literacy</p> <p>Which methods were used to collect the data? 1. Pre- and post-study survey of teachers' perceptions of classroom literacy behaviours. (p.2) 2. Parent provided information on their children via a survey (pp.6-7) 3. Letter Naming and Initial Sounds Fluency subscales of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) test 4. Total raw score of the Test of Early Reading Ability-3rd edition (TERA-3). (p.14)</p> | <p>USA</p> <p>Number of participants 86 participants</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Pre-school/nursery/kindergarten</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants Subjects were enrolled in one of four different kindergarten classes at a public elementary school in Northwest Florida. Each class was assigned one of four treatment conditions: Music/Video (sequential presentation of each condition), Music-Only, Video-Only and 'no contact' control group. (p.2) Children were from a low SES background. (p.2)</p> <p>Are outcomes reported for any of the priority groups? Low SES (intervention targeted at this group)</p> | <p>Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Arts (multi-component): Music therapy designed to teach reading skills. It involved singing, playing instruments and movement (pp.2,13)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention The study occurred between September and November of the school year (subjects had to attend a minimum of 15 treatment sessions).</p> <p>Intensity of the activity/intervention 1-2 per week (assumed): Groups lasted 25-30 minutes each (pp.9,13)</p> <p>What treatment/intervention did the control/comparison group receive? No treatment: There were four groups in total, one of which was a 'no contact' control group.</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|--|
| Schellenberg (2004) Overall weight of evidence: Medium/High | <p>What are the broad aims of the study? 'The present experiment provided a direct test of the hypothesis that music lessons enhance IQ.' (p.512)</p> <p>What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills Transferable skills Cognitive performance</p> <p>If academic achievement is measured, what subjects does the study focus on? Literacy Numeracy</p> <p>Which methods were used to collect the data? 1. Parent Rating Scale of the Behavioral Assessment System for Children (BASC) 2. Kaufman Test of Educational Achievement (K-TEA) 3. Wechsler Intelligence Scale for Children—Third Edition (WISC-III)</p> | <p>In which country/countries was the study conducted? Canada</p> <p>Number of participants 144 started (36 in each group), but 12 dropped out (keyboard: 6, voice: 4, drama: 2). Thus, analyses of pre- to post-lessons changes included data from 132 children. (p.512)</p> <p>Age of participants 6-10 years</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants The study was designed as follows: 2 intervention groups (keyboard lessons, voice lessons); 2 control groups (dance lessons, no lessons). However, in this review, drama lessons have also been compared with the 'no lessons' control group.</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Which sectors does the engagement relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: Kodaly voice lessons No: keyboard lessons, drama lessons</p> <p>Arts participation 1. Singing (<i>Kodaly voice lessons</i>) 2. Play a musical instrument (<i>keyboard lessons</i>) 3. Rehearse or perform play/drama (<i>drama lessons</i>)</p> <p>What is/are the setting(s) of the activity/intervention? Music conservatory</p> <p>Duration of the activity/intervention 36 weeks</p> <p>Intensity of the activity/intervention 1-2 per week</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: no lessons</p> |
| Standley and Hughes (1997) | <p>What are the broad aims of the study? This study evaluated the effects of music</p> | <p>In which country/countries was the study conducted?</p> | <p>Which sectors does the engagement relate to?</p> |

| Item | Study | Sample | Activity/Intervention |
|---------------------------------------|--|--|--|
| Overall weight of evidence: Medium | <p>sessions designed to enhance pre-reading and writing skills of children aged 4-5 years</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>If academic achievement is measured, what subjects does the study focus on? Literacy</p> <p>Which methods were used to collect the data? 1. A Print Awareness Test for Logos 2. A Print Concept Checklist 3. The Developmental Writing and Language Skills Checklist</p> | <p>USA</p> <p>Number of participants 32 from 2 existing classes (17, 15), after matching, 24 completed the programme (12 participants eligible in each group) (p.80)</p> <p>Age of participants 0-5 years</p> <p>Type of educational institution attended Pre-school/nursery/kindergarten</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants The pre-kindergarten Early Intervention Program in Florida is primarily targeted at four-year old children who are economically disadvantaged.</p> <p>In each group (n=12), 11 of the children were of African-American origin and one of Caucasian origin. (p.90)</p> <p>Are outcomes reported for any of the priority groups? BME, low SES, limiting disability (intervention targeted at these groups)</p> | <p>Arts</p> <p>Does the activity/intervention have a formal name? No</p> <p>Arts participation Arts (multi-component): Music lessons designed to enhance pre-reading and -writing skills of children 3-5 years. They involved movement to music, playing instruments, singing songs, drawing and illustration, counting and reading.</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 7.5 weeks</p> <p>Intensity of the activity/intervention 1-2 per week: 2 x 30 minutes sessions (a total of 15 sessions)</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: 'The control condition was instruction in the regular prekindergarten curriculum without music involvement.' (Abstract)</p> |

| Item | Study | Sample | Activity/Intervention |
|--|---|---|---|
| von Rossberg-Gempton et al. (1999) Linked report: von Rossberg-Gempton (1998) Overall weight of evidence: Medium | What are the broad aims of the study? 'The purpose of this study was to examine the potentiality of creative dance to enhance physical and cognitive functioning in seniors and young children in a rural setting.' (p.235) What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3) What learning outcomes does the study measure/report? Transferable skills: 1. <i>Social skills</i> 2. <i>Affective skills</i> Cognitive performance: 1. <i>Cognitive abilities</i> Which methods were used to collect the data? 1. The Cratty Self-Concept Scale 2. The 'happy/sad' faces scale (see linked report, pp.53, 63) 3. Questionnaire 4. WAIS-R and WISC-R intelligence tests 5. For children, the Bruininks-Oseretsky Test of Motor Proficiency was used (<i>not relevant for this review</i>) | In which country/countries was the study conducted? Canada Number of participants A total of 78 people participated, 24 older adults and 53 children. The number of original participants (before dropout) was 89. Age of participants 6-10 years 19+ years (<i>older adults – not focus of this review</i>) Type of educational institution attended Primary school Sex of participants Mixed sex Any other important features of the participants 'The children's groups were composed of heterogeneous class distribution of special needs, bright and average children.' (p.239) Are outcomes reported for any of the priority groups? No | Which sectors does the engagement relate to? Arts Does the activity/intervention have a formal name? No Arts participation Creative dance programme What is/are the setting(s) of the activity/intervention? School (in school hours) Residential home for the elderly (not focus of this review) Duration of the activity/intervention 12 weeks Intensity of the activity/intervention 30 minutes, twice a week What treatment/intervention did the control/comparison group receive? Alternative intervention: Children in a 'wait list' condition participated for 12 weeks in physical education (PE) activities such as stretching, skipping and participating in co-operative physical games. |
| Wright et al. | What are the broad aims of the study? | In which country/countries was the | Which sectors does the engagement |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|---|
| <p>(2006)</p> <p>Linked report: Wright et al. (2007)</p> <p>Overall weight of evidence: Medium/High</p> | <p>'to evaluate a community-based afterschool arts programs (combination of theatre, visual, and media arts) targeted to a low-income population. Of particular interest were whether participants would demonstrate regular and sustained attendance and whether the program would have an impact on their psychosocial functioning.' (p. 187)</p> <p>What was the design of the evaluation? Well matched comparison group study (Maryland Scale 4)</p> <p>What learning outcomes does the study measure/report? Truancy rates/behaviour problems: 1. <i>conduct</i> 2. <i>pro-social behaviour</i></p> <p>Which methods were used to collect the data? 1. National Longitudinal Survey of Children and Youth (NLSCY) 2. Teacher observations of students and rating them on 6 assessments 3. PMK (person most knowledgeable) typically the mother, completed questionnaires on socio-demographic information and scales to assess family functioning</p> | <p>study conducted? Canada</p> <p>Number of participants 183</p> <p>Age of participants 6-10 years (at baseline) The programme itself was for 9-15 year olds; participants were followed up every two years.</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants Data on the participants is taken from a longitudinal study. This survey constituted the data collection for the NAYDP (research that evaluates community-based, youth arts programmes in low-income communities). Five sites were selected for the implementation (and evaluation) of the arts programme. Therefore, the participants are from a range of 5 different sites. The population reflects an 'ethnically and culturally rich' cross-section of the Canadian population.</p> <p>Are outcomes reported for any of the priority groups? Low SES (intervention targeted at this</p> | <p>relate to? Arts</p> <p>Does the activity/intervention have a formal name? Yes: National Arts and Youth Demonstration Project (NAYDP)</p> <p>Arts participation Arts (multi-component): Combination of theatre, visual and media arts</p> <p>What is/are the setting(s) of the activity/intervention? Unclear</p> <p>Duration of the activity/intervention 6 months (and 1 day) to 1 year</p> <p>Intensity of the activity/intervention 90 minutes, twice a week</p> <p>What treatment/intervention did the control/comparison group receive? No treatment</p> |

| Item | Study | Sample | Activity/Intervention |
|------|-------|--------|-----------------------|
| | | group) | |

Sports in-depth review Appendices

Sports in-depth review: Appendix S1

References

Studies included in the in-depth review

Anderson CM (1999) The effects of martial arts training with latency age children (self-esteem, aggression, karate, sport). *Dissertation Abstracts International Section B: The Sciences and Engineering* M1–60: 1838-1838.

Freitag EJ (2006) Promoting achievement through sports: an in-depth analysis on the impact of sports and other extracurricular activities on the development of youth. *Dissertation Abstracts International Section A: Humanities and Social Sciences* M1–66: 3806-806.

Lakes KD, Hoyt WT (2004) Promoting self-regulation through school-based martial arts training. *Journal of Applied Developmental Psychology* 25(3): 283-302.

Lipscomb S (2007) Secondary school extracurricular involvement and academic achievement: a fixed effects approach. *Economics of Education Review* 26(4): 463-472.

Powell L, Gilchrist M, Stapley J (2008) A journey of self-discovery: an intervention involving massage, yoga and relaxation for children with emotional and behavioural difficulties attending primary schools. *European Journal of Special Needs Education* 23(4): 403-412.

Sharp C, Blackmore J, Kendall L, Greene K, Keys W, Macauley A, Schagen I, Yeshanew T (2003a) *Playing for Success: an evaluation of the fourth year*. Nottingham: Department for Education and Skills.

Linked studies

To Sharp et al. (2003a)

Schagen I, Kendall L, Sharp C (2002) Measuring the success of 'Playing for Success'. *Educational Research* 44(3): 255-267.

Sharp C, Chamberlain T, Morrison J, Filmer-Sankey C (2007) *Playing for Success: an evaluation of its long term impact*. DfES Research Report 844. London: DfES.

Sharp C, Kendall L, Schagen I (2003b) Different for girls? An exploration of the impact of 'Playing for Success'. *Educational Research* 45(3): 309-324.

Other references used in the text of the report

Aarnio M, Winter T, Kujala U, Kaprio J (2002) Associations of health related behaviour, social relationships, and health status with persistent physical activity and inactivity: a study of Finnish adolescent twins. *British Journal of Sports Medicine* 36: 360-364.

Chief Medical Officer (2004) *At least five a week: evidence on the impact of physical activity and its relationship to health: a report from the Chief Medical Officer*. London: Department of Health.

Coalter F, Allison M, Taylor J (2000) *The role of sport in regenerating deprived areas*. Edinburgh: The Scottish Executive Central Research Unit

- Coalter F (2001) *Realising the potential of cultural services: the case for sport*. London: Local Government Association.
- Coalter F, Allison M, Taylor J (2003) *The role of sport in regenerating deprived urban areas*. Edinburgh: Scottish Executive Central Research Unit.
- Coalter F (2004) *The benefits of sport*. Edinburgh: sportscotland.
- Coalter F (2005) *The social benefits of sport an overview to inform the community planning process*. Research Report no. 98. Edinburgh: sportscotland.
- Coe R (2002) It's the effect size, stupid: what effect size is and why it is important. Paper presented at: *British Educational Research Association annual conference*, Exeter, 12-14 September.
- Collins M, Henry I, Houlihan B, Buller J (1999) *Sport and social exclusion: a report to the Department for Culture, Media and Sport*. Loughborough: Institute of Sport and Leisure Policy, Loughborough University.
- Coopersmith S (1967) *The antecedents of self-esteem*. Palo Alto, CA: Consulting Psychologists Press.
- Daly S (2005) *Social Capital and the Cultural Sector*. London: DCMS
- DCMS (2008) *Playing to win: a new era for sport*. London: Department for Culture, Media and Sport.
- DCMS, Strategy Unit (2002) *Game plan: a strategy for delivering government's sport and physical activity objectives*. London: Cabinet Office.
- Delaney L, Keaney E (2005) *Sport and social capital in the United Kingdom: statistical evidence from national and international survey data*. London: Department for Culture, Media and Sport.
- Ekeland E, Heian F, Hagen K, Coren E (2005) Can exercise improve self esteem in children and young people? A systematic review of randomised controlled trials. *British Journal of Sports Medicine* 29(11): 792-798.
- Evans G, Shaw P (2004) *The contribution of culture to regeneration in the UK: A review of evidence*. London: DCMS
- Farrington D, Gottfredson D, Sherman L, Welsh B (2002) The Maryland Scientific Methods Scale. In: Farrington D, MacKenzie D, Sherman L, Welsh L (eds) *Evidence based crime prevention*. London: Routledge, pages 13-21.
- Foster C, Hillsdon M, Cavill N, Allender S, Cowburn G (2005) *Understanding participation in sport: a systematic review*. London: Sport England.
- Fox K (1999) The influence of physical activity on mental well-being. *Public Health Nutrition* 2: 411-418.
- Freshminds (2007) *Culture of Demand: Ways to engage a broader audience*. London: DCMS
- Galloway S (2009) Theory-based evaluation and the social impact of the arts. *Cultural Trends* 18(2): 125-148.
- Galloway S, Birkin N, Petticrew M, Hamilton C, Bell D (2006) *Quality of life and well-being: measuring the benefits of culture and sport: literature review and thinkpiece*. Edinburgh: Scottish Executive
- Goodman R (1997) The Strengths and difficulties questionnaire: A research note. *Journal of Child Psychology, Psychiatry and Allied Discipline* 40 (5): 791-9.

- Goodman R (2001) Psychometric properties of the Strengths and difficulties questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry* 40: 1337–45.
- Hurtes K, Allen L (2001) Measuring resilience in youth: The resilience attitudes and skills profile. *Therapeutic Recreation Journal* 35 (4): 333-347
- Long J, Hylton K, Spracklen K, Ratna A, Bailey S (2009) *A systematic review of the literature on black and minority ethnic communities in sport and physical recreation*. London: Sporting Equals and the Sports Councils.
- McMaster B (2008) *Supporting excellence in the arts: from measurement to judgement and recent events*. London: Department for Culture, Media and Sport.
- Policy Action Team 10 (1999) *National Strategy for Neighbourhood Renewal, Policy Action Team audit: report of the Policy Action Team 10: the contribution of sport and the arts*. London, Department for Culture, Media and Sport.
- Powell KE, Pratt M (1996) Physical activity and health. *British Medical Journal* 313: 126–127.
- Ruiz J (2004) *A literature review of the evidence base for culture, the arts and sport policy*. Edinburgh: Scottish Executive.
- Sallis J, McKenzie T, Kolody B, Lewis M, Marshall S, Rosengard P (1999) Effects of health-related physical education on academic achievement: Project SPARK. *Research Quarterly for Exercise and Sport* 70(2): 127-134.
- Scully D, Kremer J, Meade M, Graham R, Dudgeon K (1998) Physical exercise and psychological well being: a critical review. *British Journal of Sports Medicine* 32(2): 111-120.
- Sharp C, Blackmore J, Kendall L, Schagen I, Mason K, O'Connor K (2002) *Playing for Success: an evaluation of the third year*. DfES Research Report 337. London: DfES.
- Sharp C, Kendall L, Bhabra S, Schagen I, Duff J (2001) *Playing for Success: an evaluation of the second year*. DfES Research Report 291. London: DfES.
- Sharp C, Mawson C, Pocklington K, Kendall L, Morrison J (1999) *Playing for Success: an evaluation of the first year*. DfEE Research Report 167. London: DfEE.
- Shephard R (1997) Curricular physical activity and academic performance, *Pediatric Exercise Science* 9: 113-126.
- Sport England (2004) *The framework for sport in England: making England an active and successful sporting nation: a vision for 2020*. London: Sport England.
- Sport England (2008a) *Shaping places through sport*. London: Sport England.
- Sport England (2008b) *Sport England strategy 2008-2011*. London: Sport England.
- Sport Industry Research Centre (2007) *The economic importance of sport in England, 1985-2005*. Sheffield: SIRC, Sheffield Hallam University.
- Tacon R (2008) *Getting the ball rolling: sports contributions to the 2008-2011 Public Service Agreements*. London: Central Council for Physical Recreation.
- Timperley H, Wilson A, Barrar H, and Fung Y (2007) *Teacher professional learning and development: Best evidence synthesis iteration*. Wellington NZ. Ministry of Education
- UK Sport (2003) *UK strategy framework for women and sport*. London: UK Sport.
- UK Sport (2004) *The equality standard: a framework for sport*. London: UK Sport.

Sports Appendix S2:**Study quality, outcome, and effect sizes**

| Author Quality | Outcome | Group 1 N M SD | Group 2 N M SD | Effect size Hedges' g | CI lower upper |
|--|---|-------------------------------|-------------------------------|----------------------------------|-----------------------------|
| Anderson (1999) Medium | Transferable skills - social competence karate vs no sports | 41 -0.360 1.877 | 40 0.380 2.027 | -0.375 | -0.815 0.064 |
| Anderson (1999) Medium | Transferable skills - social competency baseball/ basketball vs no sports | 41 0.220 1.917 | 40 0.380 2.027 | -0.080 | -0.516 0.355 |
| Freitag (2006) Medium | Academic achievement (GPA scores): sports participation (soccer and life-skills programme) vs no sports participation | 15 0.140 0.000 | 15 -0.360 0.000 | 0.722 | -0.020 1.464 |
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (cognitive self-regulation): martial arts vs traditional physical education. Mixed Sex. | 104 0.220 0.860 | 103 - 0.260 1.090 | 0.487 | 0.211 0.764 |
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (cognitive self-regulation): martial arts vs traditional physical education. Boys only. | 104 0.350 0.720 | 103 - 0.460 1.300 | 0.769 | 0.486 1.052 |
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (cognitive self-regulation): martial arts vs traditional physical education. Girls only. | 104 0.120 0.950 | 103 - 0.090 0.820 | 0.236 | -0.038 0.509 |
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (freedom from distractibility): martial arts vs traditional physical education. Mixed sex. | 104 0.020 0.920 | 103 - 0.030 0.780 | 0.058 | -0.214 0.331 |
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (freedom from distractibility): martial arts vs traditional physical education. Boys only. | 104 0.070 0.920 | 103 0.090 0.850 | -0.022 | -0.295 0.250 |

| Author Quality | Outcome | Group 1 N M SD | Group 2 N M SD | Effect size Hedges' g | CI lower upper |
|--|---|-----------------------------|-----------------------------|--------------------------|---------------------|
| Lakes and Hoyt (2004) Medium/ High | Capacity or capability to learn/develop (freedom from distractibility): martial arts vs traditional physical education. Girls only. | 104 - 0.020 0.940 | 103 - 0.130 0.720 | 0.131 | -0.142 0.404 |
| Lakes and Hoyt (2004) Medium/ High | Academic skills (arithmetic): martial arts compared with traditional physical education. Mixed sex. | 104 0.080 0.520 | 103 - 0.090 0.490 | 0.335 | 0.061 0.610 |
| Lakes and Hoyt (2004) Medium/ High | Academic skills (arithmetic): martial arts compared with traditional physical education. Boys only. | 104 0.130 0.450 | 103 0.080 0.540 | 0.100 | -0.172 0.373 |
| Lakes and Hoyt (2004) Medium/ High | Academic skills (arithmetic): martial arts compared with traditional physical education. Girls only. | 104 0.030 0.570 | 103 - 0.100 0.460 | 0.250 | -0.024 0.523 |
| Lakes and Hoyt (2004) Medium/ High | Truancy rates/behaviour problems (conduct): martial arts vs traditional physical education. Mixed sex. | 104 0.120 0.920 | 103 - 0.210 1.090 | 0.326 | 0.052 0.600 |
| Lakes and Hoyt (2004) Medium/ High | Truancy rates/behaviour problems (conduct): martial arts vs traditional physical education. Boys only. | 104 0.070 0.890 | 103 - 0.210 1.080 | 0.282 | 0.008 0.556 |
| Lakes and Hoyt (2004) Medium/ High | Truancy rates/behaviour problems (conduct): martial arts vs traditional physical education. Girls only. | 104 0.200 0.980 | 103 - 0.210 1.160 | 0.381 | 0.106 0.656 |
| Lakes and Hoyt (2004) Medium/ High | Academic skills (arithmetic): martial arts compared with traditional physical education. | 104 0.000 0.000 | 103 0.000 0.000 | 0.418 | 0.143 0.694 |

| Author Quality | Outcome | Group 1 N M SD | Group 2 N M SD | Effect size Hedges' g | CI lower upper |
|---|---|------------------------------|--------------------------------|--------------------------|---------------------|
| Lipscomb (2007) Medium/ High | Academic achievement/ skills (numeracy) doing sport in 8 and 12 grade v no sport outcome mathematics. | 4,214 1.300 10.300 | 3,312 - 0.200 10.300 | 0.146 | 0.100 0.191 |
| Powell et al. (2008) Medium | Transferable skills (self- confidence): yoga, massage and relaxation programme vs no treatment. | 63 0.890 1.330 | 63 0.300 1.370 | 0.434 | 0.081 0.788 |
| Powell et al. (2008) Medium | Truancy rates/behaviour problems (total difficulties score): yoga, massage and relaxation programme vs no treatment. | 63 2.460 5.260 | 63 0.520 5.190 | 0.369 | 0.017 0.721 |
| Powell et al. (2008) Medium | Transferable skills (communication with teachers): yoga, massage and relaxation programme vs no treatment. | 63 0.940 1.430 | 63 0.220 1.340 | 0.516 | 0.161 0.872 |
| Powell et al. (2008) Medium | Transferable skills (concentration/attention skills): yoga, massage and relaxation programme vs no treatment. | 63 0.020 1.150 | 63 0.260 1.320 | -0.193 | -0.543 0.157 |
| Powell et al. (2008) Medium | Truancy rates/behaviour problems (conduct sub- scale): yoga, massage and relaxation programme vs no treatment. | 63 0.250 1.580 | 63 0.000 1.290 | 0.172 | -0.178 0.522 |
| Sharp et al. (2003a) Medium/ High | Academic achievement/ skills (numeracy): 'Playing for Success' sports and study skills programme vs no treatment. | 141 0.000 0.000 | 40 0.000 0.000 | 0.657 | 0.299 1.015 |
| Sharp et al. (2003a) Medium/ High | Academic achievement/ skills (reading): 'Playing for Success' sports and study skills programme vs no treatment. | 148 0.000 0.000 | 117 0.000 0.000 | 0.259 | -0.122 0.640 |

| Author Quality | Outcome | Group 1 N M SD | Group 2 N M SD | Effect size Hedges' g | CI lower upper |
|-----------------------------------|---|------------------------|------------------------|--------------------------|---------------------|
| Sharp et al. (2003a) Medium | Transferable skills (self-confidence): 'Playing for Success' sports and study skills programme vs no treatment | 162 0.000 0.000 | 90 0.000 0.000 | 0.150 | -0.068 0.367 |
| Sharp et al. (2003a) Medium | Transferable skills (independent study skills): 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 2. | 145 0.000 0.000 | 43 0.000 0.000 | 0.389 | 0.157 0.621 |
| Sharp et al. (2003a) Medium | Transferable skills (independent study skills): 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 3. | 162 0.000 0.000 | 90 0.000 0.000 | 0.269 | 0.051 0.488 |
| Sharp et al. (2003a) Medium | Academic achievement/skills (numeracy): 'Playing for Success' sports and study skills programme vs no treatment. | 125 0.000 0.000 | 91 0.000 0.000 | 0.887 | 0.604 1.170 |
| Sharp et al. (2003a) Medium | Academic achievement/skills (total ICT score) 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 2. | 171 0.000 0.000 | 31 0.000 0.000 | 1.056 | 0.660 1.453 |
| Sharp et al. (2003a) Medium | Academic achievement/skills (total ICT score) 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 3. | 192 0.000 0.000 | 118 0.000 0.000 | 0.549 | 0.315 0.782 |
| Sharp et al. (2003a) Medium | Academic achievement/skills (email) 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 2. | 171 0.000 0.000 | 31 0.000 0.000 | 0.976 | 0.582 1.371 |

| Author Quality | Outcome | Group 1 N M SD | Group 2 N M SD | Effect size Hedges' g | CI lower upper |
|--------------------------------------|---|-------------------------------|-------------------------------|----------------------------------|-----------------------------|
| Sharp et al. (2003a) Medium | Academic achievement/ skills (email) 'Playing for Success' sports and study skills programme vs no treatment. Key Stage 3. | 192 0.000 0.000 | 118 0.000 0.000 | 0.479 | 0.246 0.711 |

Sports in-depth review: Appendix S3

Structured abstracts of included studies

| Item | Study | Sample | Activity/Intervention |
|--|---|---|--|
| Anderson (1999) WoE D Medium | <p>What are the broad aims of the study? To investigate the effects of martial arts training for boys ages 8-13 on self-esteem, aggression and social competency</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills 1. <i>Social competency</i></p> <p>Which methods were used to collect the data? 1. Rosenberg self-esteem scale 2. CBCL/4-18 (Child Behaviour Checklist - includes the social competency scale and the aggressive behaviour scale)</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 41 in the karate group; 41 in the organised sports group (baseball/basketball); 40 in the comparison group</p> <p>Are outcomes reported for any of the priority groups? No</p> <p>Age of participants 6-10 years 11-15 years</p> <p>Sex of participants Male</p> <p>Any other important features of the participants Participants had recently started karate or organised basketball/baseball and had</p> | <p>Which sectors does the engagement relate to? Sport</p> <p>Sport Karate, Basketball, Baseball/Softball</p> <p>Does the activity/intervention have a formal name? No</p> <p>What is/are the setting(s) of the activity/intervention? Unclear</p> <p>Duration of the activity/intervention Unclear</p> <p>Intensity of the activity/intervention 1-2 per week</p> <p>What treatment/intervention did the control/comparison group receive? No treatment: comparison group not enrolled in karate or other organised</p> |

| Item | Study | Sample | Activity/Intervention |
|--|---|---|--|
| | | <p>had no more than two weeks of training at the time of pre-test in order to have been in the study.</p> <p>Participants were mainly white (90%) and averaging B grades or better</p> | sports (baseball or basketball) |
| Freitag (2006) WoE D Medium | <p>What are the broad aims of the study? 'The purpose of this study was to investigate the impact of the program Promoting Achievement Through Sports (PATs) on academic performance, resiliency, and school behaviour.' (p.54)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>If academic achievement is measured, what subjects does the study focus on? Not stated</p> <p>Which methods were used to collect</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants Intervention: 15 Control: 15 students</p> <p>Are outcomes reported for any of the priority groups? No</p> <p>Age of participants 11-15 years</p> <p>Type of educational institution attended Secondary school</p> | <p>Which sectors does the engagement relate to? Sport</p> <p>Sport Football (soccer)</p> <p>Does the activity/intervention have a formal name? Yes: Promoting Achievement Through Sports (PATs)</p> <p>What is/are the setting(s) of the activity/intervention? School-based extra-curricular clubs</p> <p>Duration of the activity/intervention 3 months (and 1 day) to 6 months: January to June 2004 (p.49)</p> <p>Intensity of the activity/intervention 3-4 per week: Practices were held twice a</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|---|---|
| | <p>the data?</p> <ol style="list-style-type: none"> 1. Resiliency Attitude and Skills Profile (RASP) (Hurtes and Allen, 2001) 2. Teacher's Report Form (TRF) was used to assess behaviour and other outcomes 3. School behaviour records 4. GPA data held by school | <p>Sex of participants Male</p> <p>Please specify any other important features of the participants 'All participants were of Latino descent with limited English proficiency.' (p.55)</p> | <p>week after school and games were usually played on Saturdays. (p.49)</p> <p>What treatment/intervention did the control/comparison group receive? No treatment</p> |
| <p>Lakes and Hoyt (2004)</p> <p>WoE D Medium/High</p> | <p>What are the broad aims of the study? 'To examine the utility of martial arts training for promoting self-regulation among children' (p.286)</p> <p>What was the design of the evaluation? Randomised controlled trial (Maryland Scale 5)</p> <p>What learning outcomes does the study measure/report?</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants A total of 208 were selected but the final sample included 193 students.</p> <p>Are outcomes reported for any of the priority groups? No</p> | <p>Which sectors does the engagement relate to? Sport</p> <p>Sport Taekwondo</p> <p>Does the activity/intervention have a formal name? Yes: Leadership Education Through Athletic Development (LEAD) curriculum.</p> <p>What is/are the setting(s) of the</p> |

| Item | Study | Sample | Activity/Intervention |
|------|---|--|--|
| | <p>Academic achievement/skills</p> <p>Transferable skills</p> <p>1. Affective skills (e.g. self-confidence, persistence) (RCS scale)</p> <p>2. Emotional (SDQT scale)</p> <p>3. Peer problems</p> <p>4. Pro-social behaviour (SDQT scale)</p> <p>Truancy rates/behaviour problems</p> <p>1. conduct (SDQT scale)</p> <p>Capacity or capability to learn/develop</p> <p>1. Cognitive self-regulation (e.g. ability to focus attention and efforts on task at hand) (RCS scale)</p> <p>2. Inattention/hyperactivity (SDQT scale)</p> <p>3. Freedom from distractibility (WISC-III scale)</p> <p>If academic achievement is measured, what subjects does the study focus on?</p> <p>Numeracy</p> <p>Which methods were used to collect the data?</p> | <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants Predominantly white (75%). Comparatively well off (73% family income of more than US\$100,000 per year). Between 15% and 20% of the total student body received some form of financial aid from the school.' (p.287)</p> | <p>activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 3 months (and 1 day) to 6 months: Mid-August to late January</p> <p>Intensity of the activity/intervention Other: 2-3 times per week (45 minute classes)</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual: traditional physical education</p> |

| Item | Study | Sample | Activity/Intervention |
|------|---|--------|-----------------------|
| | 1. Coopersmith (1967) Self-esteem Inventory (SEI) 2. Goodman's (1997) Strengths and Difficulties Questionnaire (SDQ). 3. Response to Challenge Scale (RCS). 4. Wechsler Intelligence Scale for Children-Third Edition (WISC-III) | | |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| Lipscomb (2007) WoE D Medium/High | <p>What are the broad aims of the study? To examine whether sport participation effects attainment and educational expectations</p> <p>What was the design of the evaluation? Well-matched comparison group study (Maryland Scale 4)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills Other</p> <p>If academic achievement is measured, what subjects does the study focus on? Numeracy Science</p> <p>Which methods were used to collect the data?</p> <p>The empirical analysis uses data from the base year through the second follow-up of the National Education Longitudinal Study of 1988.</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants 16305</p> <p>Age of participants 11-15 years</p> <p>Sex of participants Mixed sex</p> | <p>Which sectors does the engagement relate to? Sport</p> <p>Sport Basketball, Baseball/Softball, Football (soccer), Golf/pitch and putt/putting, Gymnastics, Hockey, Jogging, cross country, road running, Swimming or diving, Track and field athletics, Tennis, Volleyball</p> <p>Does the activity/intervention have a formal name? No</p> <p>What is/are the setting(s) of the activity/intervention? Not stated</p> <p>Duration of the activity/intervention Not stated</p> <p>Intensity of the activity/intervention Not stated</p> <p>What treatment/intervention did the control/comparison group receive? Alternative intervention: clubs OR sports</p> |

| Item | Study | Sample | Activity/Intervention |
|------|-------|--------|-----------------------|
| | | | activities |

| Item | Study | Sample | Activity/Intervention |
|--|--|---|--|
| Powell et al. (2008) WoE D Medium | <p>What are the broad aims of the study? 'The aim of the study reported here is to evaluate the SDP (self-discovery programme) in terms of changes in children's self-esteem, social competencies and behaviour compared to a control group.' (p.405)</p> <p>What was the design of the evaluation? Unmatched comparison group study (Maryland Scale 3)</p> <p>What learning outcomes does the study measure/report? Transferable skills</p> <p>1. <i>Self-confidence, social-confidence, communication skills, eye contact (all measured with behavioural profile tool)</i> 2. <i>Emotional, peer problems (SDQ scale)</i> 3. <i>'total difficulties' (SDQ scale)</i></p> <p>Truancy rates/behaviour problems</p> <p>1. <i>Conduct, self-control (both measured with behavioural profile tool)</i></p> | <p>In which country/countries was the study conducted? UK</p> <p>Number of participants 126 were recruited at baseline (63 in each group) but 19 were lost by follow-up. 107 participants completed both pre- and post-test measures: control (n = 54), intervention (n = 53) (pp.405-6)</p> <p>Are outcomes reported for any of the priority groups? No</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants They were children with identified</p> | <p>Which sectors does the engagement relate to? Sport</p> <p>Sport Yoga</p> <p>Does the activity/intervention have a formal name? Yes: Self-Discovery Programme (SDP)</p> <p>What is/are the setting(s) of the activity/intervention? School (in school hours)</p> <p>Duration of the activity/intervention 1 month (and 1 day) to 3 months: 12 weeks</p> <p>Intensity of the activity/intervention 1-2 per week: 45 minutes each session</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual 'The Control Group did not receive the SDP. Both the Intervention and Control Group continued to receive any additional</p> |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|--|
| | <p>2. <i>Conduct (SDQ scale)</i></p> <p>Capacity or capability to learn/develop</p> <p>1. <i>concentration, attention skills, contribution (all measured with behavioural profile tool)</i></p> <p>2. <i>hyperactivity (SDQ scale)</i></p> <p>3. <i>'Total difficulties' (SDQ scale)</i></p> <p>Which methods were used to collect the data?</p> <p>Data were collected via two sources: (1) behavioural profiles of children completed by teachers at baseline, and (2) A Strengths and Difficulties Questionnaire (SDQ: Goodman 1997, 2001) was used to supplement the behavioural profiles.</p> | <p>emotional and behavioural difficulties (and at risk of exclusion).</p> | <p>support provided.' (p.405)</p> |
| <p>Sharp et al. (2003a)</p> <p>Linked reports:</p> <p>Sharp et al. (2003b)</p> <p>Schagen et al.</p> | <p>What are the broad aims of the study?</p> <p>'The main aim of the evaluation was to provide an assessment of the effectiveness of Playing for Success and to identify and describe those features leading to success in terms of participation, gains in motivation, and positive attitudes towards learning and enhanced learning outcomes.' (p.10)</p> | <p>In which country/countries was the study conducted?</p> <p>UK</p> <p>Number of participants</p> <p>In the fourth evaluation year, 1132 pupils, 351 parents and 91 teachers took part. (p.8)</p> <p>Are outcomes reported for any of the</p> | <p>Which sectors does the engagement relate to?</p> <p>Sport</p> <p>Sport</p> <p>Basketball, Cricket, Football (soccer), Gaelic sports, Gymnastics, Hockey, Track and field athletics, Tennis, Rugby (union or league), Volleyball</p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|---|--|
| <p>(2002)</p> <p>Sharp et al. (2007)</p> <p>See also: Sharp et al. (1999, 2001, 2002)</p> <p>WoE D Medium/High (literacy and numeracy outcomes)</p> <p>Medium (ICT outcomes)</p> | <p>What was the design of the evaluation? Well-matched comparison group study (Maryland Scale 4)</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Transferable skills</p> <ol style="list-style-type: none"> 1. Independent study skills 2. Working with others <p>If academic achievement is measured, what subjects does the study focus on? Literacy Numeracy ICT</p> <p>Which methods were used to collect the data? 1. Nationally standardised tests of numeracy and reading comprehension, specially designed for the evaluation, were used to assess pupils' progress. 2. The views of pupils, parents and schools were gathered by means of</p> | <p>priority groups? No</p> <p>Age of participants 6-10 years 11-15 years</p> <p>Type of educational institution attended Primary school Secondary school</p> <p>Sex of participants Mixed sex</p> <p>Any other important features of the participants Focus is on underachieving pupils</p> | <p>Does the activity/intervention have a formal name? Yes: Playing for Success</p> <p>What is/are the setting(s) of the activity/intervention? Other sport setting: Unspecified</p> <p>Duration of the activity/intervention 1 month (and 1 day) to 3 months: The most common course length for both primary and secondary pupils is about 20 hours (over 10 weeks).</p> <p>Intensity of the activity/intervention 1-2 per week</p> <p>What treatment/intervention did the control/comparison group receive? No treatment</p> |

| Item | Study | Sample | Activity/Intervention |
|------|-----------------|--------|-----------------------|
| | questionnaires. | | |

Libraries in-depth review Appendices

Libraries in-depth review: Appendix L1

References

References for studies in the in-depth review

- Baughman J (2000) *School libraries and MCAS scores*. Boston, MA: Simmons College.
- Lance KC, Hamilton-Pennell C, Rodney MJ (2000a) *information empowered: the school librarian as an agent of academic achievement in Alaska schools. Revised edition*. Juneau, AK: Alaska State Library.
- Lance KC, Rodney MJ, Hamilton-Pennell C (2000b) *How school librarians help kids achieve standards: the second Colorado study*. Denver, CO: Colorado State Library.
- Lance KC, Rodney MJ, Hamilton-Pennell C (2000c) *Measuring up to standards: the impact of school library programs and information literacy in Pennsylvania schools*. Pennsylvania, PA: Department of Education, Office of Commonwealth Libraries.
- Lance KC, Rodney MJ and Hamilton-Pennell C (2002) *How school libraries improve outcomes for children: the New Mexico study*. San Jose, CA: Hi Willow Publishing.
- Michie JS, Chaney BW (2009) *Second evaluation of the Improving Literacy through School Libraries Program*. Jessup, MD: US Department of Education.
- Queen's University and People for Education (2006) *School libraries and student achievement in Ontario*. Toronto: Ontario Library Association.
- Rodney M, Lance K, Hamilton-Pennell C (2003) *The impact of Michigan school librarians on academic achievement: kids who have libraries succeed*. Lansing, MI: Library of Michigan.
- Shoham S (2000) Classroom collections and reading patterns. Paper presented at: *IFLA Council and General Conference*, Jerusalem, 13-18 August.
- Smith E (2001) *Texas school libraries: standards, resources, services, and students' performance*. Austin, TX: EGS Research and Consulting.
- Smith E (2006) *Student learning through Wisconsin school library media centres: teacher and student survey report*. Madison, WI: Division for Libraries, Technology and Community Learning, Wisconsin Department of Public Instruction.
- Todd RJ, Kuhlthau CC (2004) *Student learning through Ohio school libraries: background, methodology and report of findings*. Columbus, OH: Ohio Educational Library Media Association.

Other references used in the text of the report

- Arts Council England (2004) *The impact of the arts: some research evidence*. London: Arts Council England.
- Burns Owens Partnership Consulting (2009) *Capturing the impact of libraries: final report*. London: Department for Culture, Media and Sport.
- CILIP (2009) *Background information on public libraries*. <http://www.cilip.org.uk/get-involved/advocacy/public-libraries/Pages/factsaboutpubliclibraries.aspx> (accessed 21 January 2010).

- Coalter F (2001) *Realising the potential of cultural services: the case for libraries*. London: Local Government Association.
- Coalter F (2005) *The social benefits of sport: an overview to inform the community planning process*. Research Report no. 98. Edinburgh: sportscotland.
- Farrington D, Gottfredson D, Sherman L, Welsh B (2002) The Maryland Scientific Methods Scale. In: Farrington D, MacKenzie D, Sherman L, Welsh L (eds) *Evidence based crime prevention*. London: Routledge, pages 13-21.
- Galloway S (2009) Theory-based evaluation and the social impact of the arts. *Cultural Trends* 18(2): 125–148.
- Maeer G (2007) *Values and benefits of heritage: a research review*. London: Heritage Lottery Fund.
- McMaster B (2008) *Supporting excellence in the arts: from measurement to judgement and recent events*. London: Department for Culture, Media and Sport.
- MLA (2008) *Framework for the future MLA Action Plan for Public Libraries: 'towards 2013'*. London: Museums, Libraries and Archives Council.
- Naylor R, Kirkpatrick C, Oakley K, Griffin T (2006) *Review of museum, library and archive activity with children and young people: final report*. London: Museums, Libraries and Archives Council.
- Williams DW, Wavell C, Coles L (2001) *Impact of school library services on achievement and learning*. London: Museums, Archives and Libraries Council.
- Williams DW, Coles L, Wavell C (2002) *Impact of school library services on achievement and learning in primary schools: critical literature review of the impact of school library provision on achievement and learning in primary level students*. London: Museums, Libraries and Archives Council.

Libraries in-depth review: Appendix L2

Descriptive summary of studies

| Item | Study | Sample | Activity/Intervention |
|-------------------------|---|--|--|
| Baughman J (2000) | <p>Study quality Low/medium</p> <p>What are the broad aims of the study? <i>a state-wide evaluation of the presence of a school library programme (Simmons Study) and its relations on student achievement</i></p> <p>What was the design of the evaluation? (Maryland Scale 2) <i>Comparison group post-test only</i></p> <p>What learning outcomes does the study measure/report? Academic achievement/skills <i>The MCAS (Massachusetts Comprehensive Assessment System) tests</i></p> <p>Which methods were used to collect the data? Questionnaires/survey instrument Curriculum/subject-based test</p> | <p>In which county/countries was the study conducted? USA</p> <p>Number of participants 1. 1,241 elementary schools 2. 266 middle/junior high schools 3. 311 high schools</p> <p>Age of participants 6- 10 years 11-15 years 16-18 years</p> <p>Sex of participants Not stated</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? <i>Simmons Study in Massachusetts</i></p> <p>What treatment/intervention did the control/comparison group receive? Alternative interventions <i>Comparisons made between school library service with different levels of quality of library service provision</i></p> |
| Lance KC et al. (2000a) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study?</p> | <p>In which country/countries was the study conducted? USA (<i>Alaska</i>)</p> | <p>Which sectors does the engagement relate to? Libraries</p> |

| Item | Study | Sample | Activity/Intervention |
|------|---|---|--|
| | <p><i>To assess the impact of Alaska school librarians in academic achievement in the state's public schools</i></p> <p><i>To examine the direct relationship between staffing and student performance</i></p> <p><i>To identify selected activities of library media staff that affect test scores.</i></p> <p>What was the design of the evaluation? (Maryland Scale 2)</p> <p>Comparison group post-test only</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data?</p> <ol style="list-style-type: none"> 1. Questionnaire/survey instrument completed by student <i>California Achievement Test (CAT50 of reading, language, arts and mathematics)</i> 2. Questionnaire/survey instrument completed by schools/teachers <i>Questionnaire on library media policies;</i> <i>The school library's relationship to its local library</i> 3. School/college records (e.g. attendance, examination results) 4. Secondary datasets <i>Demographic/community data</i> | <p>Number of participants 211 schools <i>(105 schools with < 300 pupils;</i> <i>51 schools with 300-499 pupils;</i> <i>35 schools with 500-699 pupils;</i> <i>9 schools with 700-999 pupils;</i> <i>11 schools with ≥1000 pupils)</i></p> <p>Age of participants 6-18 years <i>1. 4th grade =157 schools (46% of sample)</i> <i>2. 8th grade (12-13 yrs)= 77 schools (23% sample)</i> <i>3. 11th grade (approx 16 yrs) = 105 schools (31% sample)</i></p> <p>Type of educational institution attended <i>1. Elementary school</i> <i>2. Middle/junior high</i></p> <p>Sex of participants Not stated</p> | <p>Does the activity/intervention have a formal name? No</p> <p>What treatment/intervention did the control/comparison group receive? <i>The analysis was based on the comparison in CAT5 scores between schools with different levels of quality of library provision</i></p> |

| Item | Study | Sample | Activity/Intervention |
|----------------------|--|---|--|
| Lance et al. (2000b) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study? <i>To confirm and update/expand the findings of the first Colorado study by measuring the impact of:</i></p> <ol style="list-style-type: none"> <i>1. academic achievement of specific leadership and collaboration activities of library media specialists;</i> <i>2. principal and teacher engagement in library media programmes;</i> <i>3. information technology, particularly networked computers offering licensed databases and the worldwide web</i> <p>What was the design of the evaluation? (Maryland Scale 2)</p> <p>Comparison group post-test only</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data?</p> <ol style="list-style-type: none"> <i>1. Questionnaire/survey instrument completed by schools</i> <i>2. School/college records</i> <i>3. Test scores - the Colorado Student</i> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants <i>124 schools (4th grad)</i> <i>76 schools (7th grade)</i></p> <p>Age of participants 6-15 years</p> <p>Type of educational institution attended Primary school Secondary school</p> <p>Sex of participants Not stated</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? <i>Information Power Guidelines</i></p> <p>What treatment/intervention did the control/comparison group receive? Not stated/not clear</p> <p>What is/are the setting(s) of the activity/intervention? Library (school/college)</p> |

| Item | Study | Sample | Activity/Intervention |
|----------------------|---|---|---|
| | <p>Assessment Program (CSAP) reading scores and cognitive performance</p> <p>4. Secondary datasets from publicly accessible websites:</p> <ul style="list-style-type: none"> - median family income - % of population below poverty line - minority % population | | |
| Lance et al. (2000c) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study? 'To build upon previous research showing a link between student academic achievement and the school library program ... which focuses on what students have learned (proficiencies) rather than what is taught (coverage of content)'</p> <p>What was the design of the evaluation? (Maryland Scale 2)</p> <p>Comparison group post-test only</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills 11th grade PSSA reading score</p> <p>Which methods were used to collect the data? 1. Questionnaires to be completed by library staff 2. Assessment completed by student:</p> | <p>In which country/countries was the study conducted? USA (Pennsylvania)</p> <p>Number of participants 435 schools (138 grade 5; 183 grade 8; 201 grade 11)</p> <p>Age of participants 6-18 years</p> <p>Type of educational institution attended 1. Elementary schools 2. Middle/junior high school 3. High/senior high school</p> <p>Sex of participants Not stated</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? <i>American Association of School Librarians new standards Information Power: Building Partnerships for Learning 1998</i></p> <p>What treatment/intervention did the control/comparison group receive? <i>No treatment</i></p> <p><i>The comparison was between the impact of library staffing levels on the 2 groups (staffed vs not staffed) on reading scores</i></p> <p>What is/are the setting(s) of the activity/intervention? Library (school/college)</p> <p>Duration of the activity/intervention Unclear</p> |

| Item | Study | Sample | Activity/Intervention |
|---------------------|---|--|---|
| | <i>Pennsylvania System of School Assessment test (PSSA reading score)</i> | | Intensity of the activity/intervention Unclear |
| Lance et al. (2002) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study? <i>To explore several issues that were explored in the other state studies, including:</i> <i>a) the identifying characteristics of school librarians and school library programmes that affect academic achievement</i> <i>b) To assess the contribution of collaboration between teachers and school librarians to the effectiveness of school library programmes</i> <i>c) To examine the growing role of information technology in school libraries, particularly licensed databases and the internet</i></p> <p>What was the design of the evaluation? (Maryland Scale 2)</p> <p>One group pre-post study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data?</p> | <p>In which country/countries was the study conducted? USA (<i>New Mexico</i>)</p> <p>Number of participants 1. 208 of 400 (52%) elementary schools serving 4th grade participated 2. 100 of the 177 (56%) junior high and middle schools serving 8th grade participated. 3. 72 of the 152 (47%) high schools serving 10th grade participated</p> <p>Age of participants 6-15 years</p> <p>Type of educational institution attended 1. Elementary schools 2. Junior high/middle schools 3. High schools</p> <p>Sex of participants Not stated</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? <i>Information Power: Building partnerships for learning</i></p> <p>What treatment/intervention did the control/comparison group receive? Not applicable (one group only)</p> <p>What is/are the setting(s) of the activity/intervention? Library (school/college)</p> <p>Duration of the activity/intervention Unclear</p> <p>Intensity of the activity/intervention Unclear</p> |

| Item | Study | Sample | Activity/Intervention |
|--------------------------|---|---|--|
| | <p>1. Questionnaire/survey instrument completed by schools/teachers:</p> <p>2. Secondary datasets: The New Mexico Department of Education for data on number and % of students eligible for free/subsidised meals; ethnicity; school expenditure; teachers' characteristics</p> <p>3. Educational attainment: test scores (New Mexico Achievement Assessment program)</p> | | |
| Michie and Chaney (2009) | <p>Study quality Medium</p> <p>What are the broad aims of the study? Several aims, the relevant one: How do reading achievement scores vary in schools that received grants for 1, 2 or 3 years compared with matched comparison schools that have not received grants?</p> <p>What was the design of the evaluation? (Maryland Scale 3) Unmatched comparison group study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data? Mean Test Scores in Reading</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants The survey of school libraries was sent to 800 schools: School libraries receiving grants (grantees): N= 400 matched sample of school libraries eligible for the grant in that year (non-grantees): N= 400 Of the 701 respondents, test scores were available for 553 schools (79%). Baseline data were available for only 40% of the sample.</p> <p>Age of participants Unclear</p> <p>Type of educational institution</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? Improving Literacy Through School Libraries (LSL) program</p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual (received no grants) The comparison was between schools with grants and schools with no grants</p> <p>What is/are the setting(s) of the activity/intervention? Library (school/college)</p> <p>Duration of the activity/intervention Grants for the LSL program are for one year</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|--|
| | | attended Secondary and primary Sex of participants Mixed sex | |
| Queen's University and People for Education (2006) | Study quality Low/Medium What are the broad aims of the study? <i>To analyse the relationship between student achievement (test scores and attitudes toward reading) and library resources and staff</i> What was the design of the evaluation? (Maryland Scale 2) Comparison group post-test only What learning outcomes does the study measure/report? Academic achievement/skills Which methods were used to collect the data? 1. Questionnaire/survey instrument completed by teacher 2. Curriculum/subject-based test, completed by student | In which country/countries was the study conducted? Canada Number of participants <i>Includes assessment data from more than 50,000 students and over 800 of Ontario's publicly-funded elementary schools</i> <i>The study examined data for 26,643 grade 3 students from 707 schools and 27,565 grade 6 students from 668 schools.</i> Age of participants 6-10 years Type of educational institution attended <i>Elementary schools (Grades 3-6)</i> Sex of participants Not stated | Which sectors does the engagement relate to? Libraries Does the activity/intervention have a formal name? No What treatment/intervention did the control/comparison group receive? <i>Unclear (no details provided)</i> What is/are the setting(s) of the activity/intervention? <i>School libraries</i> Duration of the activity/intervention Not stated Intensity of the activity/intervention Not stated |

| Item | Study | Sample | Activity/Intervention |
|----------------------|--|---|---|
| Rodney et al. (2003) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study? <i>To assess the impact of Michigan School librarians on academic achievement</i></p> <p>What was the design of the evaluation? (Maryland Scale 2) One group pre-post study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills <i>Scores on the reading portion of the standards-based Michigan Educational Assessment Program (MEAP)</i></p> <p>Which methods were used to collect the data? 1. Questionnaire/survey instrument 2. Secondary datasets <i>Educational attainment data was extracted from the United States Census Bureau Web site.</i> <i>The test scores used are 2001 scores on the reading portion of the standards-based Michigan Educational Assessment Program (MEAP).</i></p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants Not reported</p> <p>Age of participants 6-18 years (assumed) <i>Grades 4, 7 and 11</i></p> <p>Type of educational institution attended Primary school Secondary school</p> <p>Sex of participants Mixed sex</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? <i>No</i></p> <p>What treatment/intervention did the control/comparison group receive? Not applicable (one group only)</p> <p>What is/are the setting(s) of the activity/intervention? <i>School library</i></p> <p>Duration of the activity/intervention Unclear</p> <p>Intensity of the activity/intervention <i>Continuous</i></p> |
| Shoham | Study quality | In which country/countries was the | Which sectors does the engagement |

| Item | Study | Sample | Activity/Intervention |
|--------------|---|---|---|
| (2000) | <p>Low/Medium</p> <p>What are the broad aims of the study? <i>To investigate the relationship between the type of school library and the pupils' reading habits over the long term. The hypothesis was that pupils who have a central library in their school together with classroom collections read more than pupils whose school only had a central library or only classroom collections</i></p> <p>What was the design of the evaluation? (Maryland Scale 2) Comparison group post-test only</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills Reading</p> <p>Which methods were used to collect the data? <i>Questionnaires were distributed among a sample of 300 pupils</i></p> | <p>study conducted? <i>Israel</i></p> <p>Number of participants 301 elementary school children from 4 schools: group A (with school library) n=107; group B (no school library but with classroom collection) n=91; group C (no participation) n=103</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Elementary schools</p> <p>Sex of participants Not stated</p> <p>Please specify any other important features of the participants <i>The socio-demographic characteristics of the four schools' pupils were quite similar: centre of the country, urban areas, middle-class</i></p> | <p>relate to? Libraries</p> <p>Does the activity/intervention have a formal name? No</p> <p>What treatment/intervention did the control/comparison group receive? 1. Group B: children who participated in the class library in grades 2-3, with no central library in the school, but with classroom collection in all the classes including the fourth grade. 2. Group C: children who did not participate in a class-library project, while having a central library in the school.</p> <p>What is/are the setting(s) of the activity/intervention? School library</p> <p>Duration of the activity/intervention 2 academic years</p> <p>Intensity of the activity/intervention Not stated</p> |
| Smith (2001) | <p>Study quality Low/Medium</p> <p>What are the broad aims of the study? 3 main aims, one relevant:</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants</p> | <p>Which sectors does the engagement relate to? School library</p> <p>Does the activity/intervention have a</p> |

| Item | Study | Sample | Activity/Intervention |
|--------------|--|--|--|
| | <p><i>To determine the impact that school libraries have on student performance</i></p> <p>What was the design of the evaluation? (Maryland Scale 2) comparison group post-test only</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data? Curriculum/subject-based test (assessment/exam etc. completed by student) <i>The Texas Assessment of Academic Skills (TAAS)</i></p> | <p><i>Original sample of schools (with librarians) approached: N=600 (327 Elementary schools; 120 Middle/junior high school; 139 High schools; 14 Elementary-secondary schools).</i></p> <p><i>The analyses were performed on 1,779 schools without librarians</i></p> <p><i>Sample size used in the analyses:</i> <i>All schools with librarians (N=500)</i> <i>All schools without librarians (N=1,541)</i></p> <p>Age of participants 6-18 years</p> <p>Type of educational institution attended <i>Elementary schools</i> <i>Middle/Junior high schools</i> <i>High schools</i></p> <p>Sex of participants Not stated</p> | <p>formal name? <i>The Texas Study</i></p> <p>What treatment/intervention did the control/comparison group receive? <i>schools with no librarian service</i></p> <p>What is/are the setting(s) of the activity/intervention? School</p> |
| Smith (2006) | <p>Study quality Low</p> <p>What are the broad aims of the study? <i>To understand the ways and extent to which students benefit from school libraries</i></p> <p>What was the design of the evaluation?</p> | <p>In which country/countries was the study conducted? USA</p> <p>Number of participants <i>107 teachers, 3,957 students, 51 schools</i></p> <p>Age of participants</p> | <p>Which sectors does the engagement relate to? Libraries</p> <p>Does the activity/intervention have a formal name? No</p> |

| Item | Study | Sample | Activity/Intervention |
|--------------------------|---|--|---|
| | <p>(Maryland Scale 1)</p> <p>Single-group post-test study</p> <p>What learning outcomes does the study measure/report?</p> <p>Academic achievement/skills</p> <p>Transferable skills</p> <p>Capacity or capability to learn/develop</p> <p><i>Students asked whether libraries helped their</i></p> <p>Which methods were used to collect the data?</p> <p>Questionnaire/survey instrument completed by student</p> <p>Questionnaire/survey instrument completed by teacher</p> | <p>6-18 years</p> <p>Type of educational institution attended</p> <p>Primary schools</p> <p>Secondary schools</p> <p>Sex of participants</p> <p>Mixed sex</p> | <p>What treatment/intervention did the control/comparison group receive?</p> <p>One group only</p> <p>What is/are the setting(s) of the activity/intervention?</p> <p>School library</p> <p>Duration of the activity/intervention</p> <p><i>School libraries always available, not clear how accessible they are to students, or how much students used them</i></p> <p>Intensity of the activity/intervention</p> <p>not clear</p> |
| Todd and Kuhlthau (2004) | <p>Study quality</p> <p>Low</p> <p>What are the broad aims of the study?</p> <p><i>To assess the help (impact) of high quality school libraries in Ohio on students</i></p> <p>What was the design of the evaluation?</p> <p>(Maryland Scale 1)</p> <p>Single-group post-test study</p> <p>What learning outcomes does the study measure/report?</p> <p>Academic achievement/skills</p> <p>Transferable skills</p> | <p>In which country/countries was the study conducted?</p> <p>USA</p> <p>Number of participants</p> <p><i>Not clear (between 10,000 and 13,000 responses)</i></p> <p>Age of participants</p> <p>6-18 years</p> <p>Type of educational institution attended</p> <p>Primary schools</p> <p>Secondary schools</p> | <p>Which sectors does the engagement relate to?</p> <p>Libraries</p> <p>Does the activity/intervention have a formal name?</p> <p>No</p> <p>What treatment/intervention did the control/comparison group receive?</p> <p>One group only</p> <p>What is/are the setting(s) of the activity/intervention?</p> <p>Library</p> |

| Item | Study | Sample | Activity/Intervention |
|------|---|---|--|
| | Personal development Attitude to learning Capacity or capability to learn/develop Motivation for learning Other: <i>students asked whether library helps them complete their school work</i> Which methods were used to collect the data? Questionnaire/survey instrument completed by student | Sex of participants Mixed sex | Duration of the activity/intervention <i>School libraries provided continuously, but unclear how much access students had or how much they used them</i> Intensity of the activity/intervention Unclear |

Museum, galleries and heritage in-depth review appendices

Museums, galleries and heritage in-depth review: Appendix M1

References

References for studies in the in-depth review

AEA Consulting (2005) Tyne & Museums Bristol Museums, Galleries and Archives. Social Impact Programme Assessment. Newcastle. Tyne & Wear Museums.

Hooper-Greenhill E, Dodd J, Philips M, Jones C, Woodham J, O'Riain H (2004a) Inspiration, identity, learning: the value of museums. The evaluation of DCMS/DfES Strategic Commissioning 2003-4: National/Regional Museum Education Partnerships . Leicester: Research Centre for Museums and Galleries.

Hooper-Greenhill E, Dodd J, Phillips M, O'Riain H, Jones C, Woodward J (2004b) What Did You Learn at the Museum Today? The evaluation of the impact of the Renaissance in the Regions Education Programme in the three Phase 1 Hubs (August, September and October 2003). Leicester: Research Centre for Museums and Galleries.

Hooper-Greenhill E, Dodd J, Gibson L, Phillips M, Jones C, Sullivan E (2006) What Did You Learn at the Museum Today? Second Study. Evaluation of the outcomes and impact of learning through the implementation of the Education Programme Delivery Plan across nine Regional Hubs (2005). Leicester: Research Centre for Museums and Galleries.

Hooper-Greenhill E, Dodd J, Creaser C, Sandell R, Jones C, Woodward A (2007) Inspiration, identity, learning: the value of museums second study. An evaluation of the DCMS/DCSF National/Regional Museum Partnership Programme in 2006/2007. London: Department for Culture, Media and Sport .

Kanevsky L, Corke M, Frangkiser L (2008) The Academic Resilience and Psychosocial Characteristics of Inner-City English Learners Participating in a Museum-Based School Program. *Education and Urban Society* 40(4): 452 - 475.

Melber LM (2003) Partnerships in Science Learning: Museum Outreach and Elementary Gifted Education. *Gifted Child Quarterly* 47(4): 251 - 258.

Randi Korn Associates (2007) Teaching literacy through art. Final report: synthesis of 2004-05 and 2005-06 studies.. New York: Solomon R. Guggenheim Museum.

Stanley J (2008) MAGPIE Phase 3 Final Report: Progress in achievement. Coventry: Centre for Education and Industry.

Watson S, Dodd J, Jones C (2007) The impact of museum visits on the attainment of secondary pupils in the East of England 2006-2007. Leicester: Research Centre for Museums and Galleries.

West H (2008) A study of the impact of using the TASC wheel on children's involvement in their own learning. *Gifted Education International* 24(2-3): 297 - 304.

Other references used in the text of the final report

Arts Council England (2004) *The impact of the arts: some research evidence*. London. Arts Council England.

Burns Owens Partnership (2005) *New directions in social policy: developing the evidence base for museums, libraries and archives in England*. London: Museums, Libraries and Archives Council.

Coalter FR (2005) *The Social Benefits of Sport An Overview to Inform the Community Planning Process*. Edinburgh. Sportscotland. Research Report no. 98

DCMS (1999) *Museums for the Many: Standards for Museums and Galleries to Use When Developing Access Policies*. London: Department for Culture, Media and Sport

DCMS (2000a) *Centres for Social Change: Museums, Galleries and Archives for All*. London: Department for Culture, Media and Sport

DCMS (2000b) *The Learning Power of Museums: A Vision for Museum Education*. London: Department for Culture, Media and Sport

DCMS (2010) *The Government's statement on the historic environment for England 2010*. London. Department for Culture Media and Sport.

Desai, P. and Thomas, A (1998) *Cultural Diversity: Attitudes of Ethnic Minority Populations Towards Museums and Galleries*. Museums and Galleries Commission: London.

English Heritage (2000) *Power of place: The future of the historic environment*. London. English Heritage

Farrington D, Gottfredson D, Sherman L, Welsh B (2002) *The Maryland Scientific Methods Scale In* (eds) Farrington D, MacKenzie D, Sherman L, Welsh L.. *Evidence Based Crime Prevention*. London. Routledge. Pp 13-21

Galloway S (2009) *Theory-based evaluation and the social impact of the arts* *Cultural Trends* Vol. 18, No. 2, June 2009, 125–148

Hooper- Greenhill, E (1998) 'Audiences: A curatorial dilemma', in *The Educational Role of the Museum*, ed. E. HOOPER-GREENHILL, .London. Routledge,

Maeer G (2007) *Values and benefits of heritage: A research review*. London Heritage Lottery fund

McMaster B (2008) *Supporting excellence in the arts - from measurement to judgement and recent events*. London: Department for Culture, Media and Sport

MLA (2008) *Framework for the Future MLA Action Plan for Public Libraries – "towards 2013"* London. Museums, Libraries and Archives Council.

MORI (1999) The Popularity of UK Museums. Report to the Museums and Galleries Commission. London. Ipsos Mori.

Naylor R, Kirkpatrick C, Oakley K, Griffin T (2006) .Review of Museum, Library and Archive Activity with Children and Young People: Final Report. London. Museums Libraries and Archives Council

Tlili, A, Gewirtz, S. and Cribb, A (2007) 'New Labour's Socially Responsible Museum', *Policy Studies*, 28: 3, 269 — 289

Travers T (2006) Museums and Galleries in Britain Economic, social and creative impacts. London. Museums Libraries and Archives Council

Wavell C, Baxter G, Johnson I, Williams D (2002) Impact evaluation of museums archives and libraries: Available evidence project. London. Museums Libraries and Archives Council

Museums, Libraries and Heritage in-depth review: Appendix M2:
Descriptive summary of studies

| Item | Study | Sample | Activity/Intervention |
|--|---|--|--|
| AEA Consulting (2005) Weight of evidence D: Overall weight of evidence Low | What are the broad aims of the study? <i>To assess the impact of the Rwanda Exhibition Project on young people's cultural understanding, creative and artistic learning</i> What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study What learning outcomes does the study measure/report? Personal development <i>self confidence</i> Cultural knowledge Motivation for learning Creativity Other <i>Learn something new</i> Which methods were used to collect the data? <i>Survey instrument (not known who completed)</i> | In which country/countries was the study conducted? UK Number of participants 12 Age of participants 6-10 years Type of educational institution attended Primary school Sex of participants Not stated | Which sectors does the engagement relate to? Arts and Museums Does the activity/intervention have a formal name? <i>The Rwanda Educational Exhibition Project</i> What treatment/intervention did the control/comparison group receive? Not applicable (one group only) What is/are the setting(s) of the activity/intervention? <i>Arts Museum in Bristol</i> Duration of the activity/intervention <i>4 months</i> Intensity of the activity/intervention <i>Not stated</i> |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|---|
| <p>Hooper-Greenhill E et al. (2004a)</p> <p>Weight of evidence D: Overall weight of evidence Low</p> | <p>What are the broad aims of the study? <i>To explore the impact of museum activities on learning for schools and community groups</i></p> <p>What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills <i>Knowledge and understanding (One of the 5 Generic Learning Outcomes [GLO] identified)</i> Transferable skills <i>Skills (One of the 5 GLO identified)</i> Cultural knowledge <i>Action, behaviour, progression (One of the 5 GLO identified)</i> Curiosity/Motivation for learning <i>Attitudes and values/Enjoyment, inspiration, creativity (Two of the 5 GLO identified)</i></p> <p>Which methods were used to collect the</p> | <p>In which county/countries was the study conducted? UK</p> <p>Number of participants <i>29,701 contacts with school pupils, 9415 completed questionnaires: 79% responses from Key Stage 2 (n=7438; 21% responses from Key Stage 3 (n=1977)</i></p> <p><i>503 teachers completed questionnaires</i></p> <p>Are outcomes reported for any of these priority groups? No</p> <p>Age of participants 6-10 years 11-15 years 16-18 years</p> <p>Type of educational institution attended Primary schools <i>71% primary schools/nursery</i> Secondary schools <i>18% secondary schools/colleges</i></p> | <p>Which sectors does the engagement relate to? museums</p> <p>Does the activity/intervention have a formal name? <i>2003-2004: National/Regional Museum Education Partnerships</i></p> <p>What treatment/intervention did the control/comparison group receive? Not applicable (one group only)</p> <p>What is/are the setting(s) of the activity/intervention? Museum (art/history and science) <i>Including: National Gallery, National Portraits Gallery, British Museum, Tate Britain, V&A; National Museum of Science and Industry</i></p> <p>Duration of the activity/intervention Not stated</p> <p>Intensity of the activity/intervention Not stated</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|---|
| | data? Questionnaire/survey instrument <i>completed by student</i> Focus group interview <i>interviews, focus groups</i> Other <i>Case studies, site visits</i> | Other <i>7% middle, special and independent schools</i> Sex of participants Mixed sex Please specify any other important features of the participants <i>1. A significant 30% of schools from areas of deprivation where child poverty is high'</i> <i>2. Schools from rural and urban regions</i> | |
| Hooper-Greenhill E et al. (2004b) Weight of evidence D: Overall weight of evidence Low | What are the broad aims of the study? <i>Several aims, relevant ones:</i> <i>1. Identify the learning that has taken place and analyse this against the generic learning outcomes (GLO) and the outcomes posed by the DfES</i> <i>2. Relate the teacher's perceptions of their learning to the perceptions of the pupils themselves</i> What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study | In which county/countries was the study conducted? UK Number of participants <i>17198 KS2 pupils</i> <i>3406 KS3 pupils</i> <i>About 1,000 teachers completed questionnaires;</i> <i>71 attended focus-group workshops</i> Are outcomes reported for any of | Which sectors does the engagement relate to? Museums Does the activity/intervention have a formal name? <i>Renaissance in the Regions</i> What treatment/intervention did the control/comparison group receive? Not applicable (one group only) What is/are the setting(s) of the |

| Item | Study | Sample | Activity/Intervention |
|------|---|--|---|
| | <p>What learning outcomes does the study measure/report? Transferable skills <i>Activity, behaviour, progression (One of the 5 GLO identified)</i> Cultural knowledge <i>Skills/knowledge and understanding (One of the 5 GLO identified)</i> Attitude to learning <i>Attitudes and values (One of the 5 GLO identified)</i> Creativity <i>Enjoyment, inspiration and creativity (One of the 5 GLO identified)</i></p> <p>Which methods were used to collect the data? Questionnaire/survey instrument 1. <i>Post-visit questionnaires by pupils</i> 2. <i>Post-visit questionnaires by teachers</i> Focus group interview of teachers School/college records 1. <i>Numerical data collection of pupils usage 2002, 2003 - template for museums to record information about school visits</i> 2. <i>Activities for school-age children during summer holidays 2003 - Template for museums to record information about the range of holiday activities and number of</i></p> | <p>these priority groups? No</p> <p>Age of participants 6-10 years 11-15 years</p> <p>Type of educational institution attended Primary school 78% secondary schools 13%</p> <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants 1. <i>Over 28% of the visits from schools located in wards as being amongst the 10% most deprived wards in England</i> 2. <i>465 visits made by schools located in wards which fell into the 20% most deprived ward in England.</i></p> | <p>activity/intervention? Museums (arts/history/science) <i>Including:</i> <i>Bristol Museums and art gallery;</i> <i>Plymouth City Museum and Art Gallery;</i> <i>Royal Cornwall Museum, Truro; Exeter City Museums and Art Gallery; the North of England Open Air Museum;</i> <i>Birmingham Museums and Art gallery;</i> <i>Coventry Arts and Heritage; Potteries Museums and Art Gallery;</i> <i>Wolverhampton Arts and Museums</i></p> <p>Duration of the activity/intervention One day or less</p> <p>Intensity of the activity/intervention Once <i>one visit</i></p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|---|---|
| | <i>participants</i> | | |
| Hooper-Greenhill E et al. (2006) Weight of evidence D: Overall weight of evidence Low | What are the broad aims of the study? <i>To explore the impact of Renaissance funding for museum education, looking specifically at the impact on learning following school visits to museum</i> What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study What learning outcomes does the study measure/report? Transferable skills <i>Skills (One of the 5 Generic Learning Outcomes [GLO] identified)</i> Cultural knowledge <i>Knowledge and understanding (One of the 5 GLO identified)</i> Attitude to learning <i>Attitudes and values (One of the 5 GLO identified)</i> Capacity or capability to learn/develop <i>Action, behaviour, progression (One of the 5 GLO identified)</i> Curiosity/motivation/creativity/enjoyment <i>enjoyment, inspiration, creativity (One of the 5 GLO identified)</i> | In which county/countries was the study conducted? UK Number of participants <i>26,791 pupils of all ages completed questionnaires</i> <i>82% from ≤KS2 (N= 21,845)</i> <i>18% from ≥KS3 (N=4,946)</i> Are outcomes reported for any of these priority groups? No Age of participants 6-10 years 11-15 years 16-18 years Type of educational institution attended Primary school <i>Key Stage 2 and below</i> Secondary school <i>Key Stage 3 and above</i> Post-16 education <i>6th form, FE college</i> | Which sectors does the engagement relate to? Museums Does the activity/intervention have a formal name? <i>Education Programme Delivery Plan across 9 Regional Hubs (Renaissance)</i> What treatment/intervention did the control/comparison group receive? Not applicable (one group only) What is/are the setting(s) of the activity/intervention? Museum (Arts, History/Natural History) including: <i>The Collection, Lincoln;</i> <i>Manchester Art Gallery; Tullie House Museum and Art Gallery Carlisle;</i> <i>Lincolnshire Museums Service;</i> <i>Norfolk Museums and Archaeology Service; Colchester Museums;</i> <i>Luton Museums Service; Museum in Docklands; Bolton Museum;</i> <i>Hampshire Museums and Archives Service; Armley Mills</i> |

| Item | Study | Sample | Activity/Intervention |
|--|--|---|---|
| | <p>Which methods were used to collect the data? Questionnaire/survey instrument completed by student and teachers Focus group interview teachers Other 1. Three in-depth school case studies 2. two-day long seminars with museum staff</p> | <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants 1. 32% of school visits made by schools located in areas of deprivation 2. 12% of the schools were special schools</p> | <p><i>Also visited/participated: Museums of Phase 1 (2003 study), which included 69 museums nationwide (Arts, History and natural history museums)</i></p> <p>Duration of the activity/intervention 1 month (and 1 day) to 3 months <i>School visits were made to the museums in September and October 2005</i></p> <p>Intensity of the activity/intervention Once <i>There were 1594 single school visits to museums</i></p> |
| <p>Hooper-Greenhill E et al. (2007)</p> <p>Weight of evidence D: Overall weight of evidence Low</p> | <p>What are the broad aims of the study? <i>Several aims, relevant ones:</i> 1. To provide a quantitative measure of who benefited from the projects 2. To identify the learning outcomes of the education projects for teachers and pupils, quantitatively (using the Generic Learning Outcomes [GLO]) and qualitatively 3. To review the social value of the museums' education and community work</p> <p>What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study</p> <p>What learning outcomes does the study measure/report?</p> | <p>In which county/countries was the study conducted? UK</p> <p>Number of participants 47,511 pupil contacts, 7,253 pupils completed questionnaires (3,916 from KS2 and below, 3,337 from KS3 and above)</p> <p>Are outcomes reported for any of these priority groups? Women <i>pupils' views were also presented by gender</i></p> <p>Age of participants</p> | <p>Which sectors does the engagement relate to? Museums</p> <p>Does the activity/intervention have a formal name? <i>The DCMS/DCSF National/Regional Museum Partnership Programme 2006-2007</i></p> <p>What treatment/intervention did the control/comparison group receive? Not applicable (one group only)</p> <p>What is/are the setting(s) of the activity/intervention? 1. Museum (art, science, history and</p> |

| Item | Study | Sample | Activity/Intervention |
|--|--|--|---|
| | <p>Transferable skills <i>Skills (one of the GLO identified)</i> Cultural knowledge <i>Knowledge and understanding (one of the GLO identified)</i> Creativity/motivation <i>Enjoyment, Inspiration, Creativity (one of the GLO identified)</i></p> <p>Which methods were used to collect the data? Questionnaire/survey instrument 1. Completed by student 2. Completed by teachers, community group leaders, museum and participants in community events</p> | <p>6-10 years 11-15 years 16-18 years</p> <p>Type of educational institution attended Primary school Key Stage 2 Secondary school Key Stage 3,4,5 post-16 education</p> <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants <i>the schools taking part are located in areas which experience a range of deprivation, with no clear preponderance of schools from areas of high or low social deprivation</i></p> | <p>natural history) <i>Including: The National Gallery; Nottingham City Museums and Galleries; Cartwright Hall Museum and Art Gallery; Victoria and Albert Museum; Birmingham Museum and Art Gallery; The National Museum of Science and Industry; National Media Museum, The Science Museum; The Natural History Museum; Oxford University Museum of Natural History; The National Maritime Museum; National Museums Liverpool; Bristol City Museums etc</i></p> <p>Duration of the activity/intervention Not stated</p> <p>Intensity of the activity/intervention Not stated</p> |
| <p>Kanevsky L et al. (2008) The</p> <p>Weight of evidence D: Overall weight of evidence Medium</p> | <p>What are the broad aims of the study? <i>To track academic and psychosocial aspects of resilience development in high-poverty, inner-city elementary school students not yet proficient in English during</i></p> | <p>In which county/countries was the study conducted? USA</p> <p>Number of participants.</p> | <p>Which sectors does the engagement relate to? Museums</p> <p>Does the activity/intervention have a</p> |

| Item | Study | Sample | Activity/Intervention |
|------|--|---|--|
| | <p>2 years of participation in an enriched, museum-based intervention (<i>School in the Park [SITP]</i>)</p> <p>What was the design of the evaluation? (Maryland Scale 3) Unmatched comparison group study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills 1. <i>academic resilience</i> Transferable skills 1. <i>academic self-concept</i> 2. <i>academic self-efficacy</i> Attitude to learning 1. <i>attitude toward school</i></p> <p>Which methods were used to collect the data? Questionnaire/survey instrument <i>completed by students</i> Curriculum/subject-based test <i>completed by students</i> 1. SAT9, the nationally normed math achievement test annually administered to all students in Grades 2 through 11 in California public schools. 2. The CMST, a criterion-referenced test designed to assess students' performance on California's academic content standards</p> | <p>1. All SITP participants (N=553) (268 in Grade 3, 285 in Grade 4). 2. A comparison group (N=301) (139 in Grade 3, 162 in Grade 4). 3. Complete data collected and analyzed from 171 SITP participants (114 Grade 3-4; 57 Grade 4-5) 4. There was 30 students from the comparison school (11 Grade 3-4 and 19 Grade 4-5)</p> <p>Are outcomes reported for any of these priority groups? No</p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary school <i>Elementary School</i></p> <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants 1. Both schools located in the same economically and academically</p> | <p>formal name? <i>School in the Park (SITP)</i></p> <p>What treatment/intervention did the control/comparison group receive? Treatment as usual (<i>without exposure to SITP</i>) <i>The comparison was between schools with exposure to SITP vs comparison group (without exposure to SITP)</i></p> <p>What is/are the setting(s) of the activity/intervention? Museum (arts/science/history/natural history) <i>Including:</i> <i>Museum of Art; Museum of Photographic Arts; Science Centre; Museum of Natural History; Historical Society; Museum of Man; Hall of Champions; Aerospace Museum; Arts setting</i> <i>Junior Theatre</i> <i>Other</i> <i>Zoo</i></p> <p>Duration of the activity/intervention 1 month (and 1 day) to 3 months 1. <i>Students worked with museum educators for up to 2.5 hours in the</i></p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| | Secondary datasets <i>Demographic, attendance, and achievement test data were provided by the Standards, Assessment and Accountability Division of San Diego</i> | <i>challenged community. Nearly all students qualify for the federally subsidized lunch program, and 75% to 80% of the students are not fluent in English</i> <i>2. inner-city school in a predominantly Hispanic, low-income, high-crime neighborhood</i> | <i>morning. For the remainder of the day, they worked with their regular classroom teacher on regular core curriculum</i> <i>2. During the years, 3rd graders attended SITP for a total of 9 weeks; 4th graders attended 7 weeks; 5th graders attended 5 weeks</i> Intensity of the activity/intervention 5-7 days per week 5 days (Monday to Friday) |
| Melber LM (2003) Weight of evidence D: Overall weight of evidence Low | What are the broad aims of the study? <i>To investigate the impact of a specially designed museum science program on academically gifted elementary students on changes in</i> <i>a) attitude toward science careers,</i> <i>b) understanding of scientific work and scientists</i> <i>c) content knowledge gains of participants.</i> What was the design of the evaluation? (Maryland Scale 2) One group pre-post study What learning outcomes does the study measure/report? Academic achievement/skills <i>content knowledge gains of participants</i> Other <i>1. changes in attitude toward science</i> | In which county/countries was the study conducted? USA Number of participants <i>31 (4th and 5th grade students)</i> Age of participants 6-10 years Type of educational institution attended <i>Elementary school</i> Sex of participants <i>Mixed sex</i> Please specify any other important features of the participants | Which sectors does the engagement relate to? Museums Does the activity/intervention have a formal name? No What treatment/intervention did the control/comparison group receive? Not applicable (one group only) What is/are the setting(s) of the activity/intervention? <i>1. School-based extra-curricular clubs museum outreach programme</i> <i>"Eight of the sessions took place at the school site using museum specimens and artifacts.... "</i> |

| Item | Study | Sample | Activity/Intervention |
|---|---|---|---|
| | <p>careers</p> <p>2. <i>changes in understanding of scientific work and scientists</i></p> <p>Which methods were used to collect the data?</p> <p>1. Questionnaire/survey instrument completed by student <i>Data were collected on students' views of scientists and what scientists do, a version of the Draw-a-Scientist Test (DAST; Chambers, 1983) was used</i></p> <p>2. Questionnaire/survey instrument completed by parent <i>Parental post-questionnaires to indicate what aspect of the course they felt was most beneficial to their child.</i></p> | <p>1. <i>The schools were in a middle/upper-middle-class suburban neighborhood of Los Angeles County.</i></p> <p>2. <i>The school consisted of a high-achieving student body overall, scoring 9 (out of 9) on an Academic Performance Index scale</i></p> <p>3. <i>the ethnic background of the class: 13% Asian/Asian American, 6% African American, 81% Caucasian, similar to schoolwide representation of these three ethnic groups: 7%, 7%, and 70%, respectively</i></p> | <p>2. Museum (natural history) urban museum of natural history <i>"The ninth session took place at the museum, where students were able to meet scientists and visit the working areas of the museum that are normally closed to the general public..., the students participated in inquiry-based activities, explored the halls and critically examined objects on display, and interacted with museum scientists and the collections."</i></p> <p>Duration of the activity/intervention <i>Approximately one month (8 school-based sessions - held twice weekly - plus one trip to the museum).</i></p> <p>Intensity of the activity/intervention <i>twice a week</i></p> |
| <p>Randi Korn Associates (2007)</p> <p>Weight of evidence D: Overall weight of evidence Low/Medium</p> | <p>What are the broad aims of the study? <i>To measure two sets of LTA (Literacy Through Arts) program outcomes: the teacher, teaching artist, and student outcomes of the program and literacy-related teacher and student outcomes related to the New York State English Language Arts Learning Standards</i></p> <p>What was the design of the evaluation? (Maryland Scale 4) Well matched comparison group study</p> | <p>In which county/countries was the study conducted? USA</p> <p>Number of participants 605 third graders who completed questionnaires: 1. 215 from Treatment Group A (LTA program only), 2. 190 from Treatment Group B (LTA program and teacher professional development),</p> | <p>Which sectors does the engagement relate to? Museums</p> <p>Does the activity/intervention have a formal name? <i>Literacy Through Arts</i></p> <p>What treatment/intervention did the control/comparison group receive? No treatment <i>The Control Group was defined as</i></p> |

| Item | Study | Sample | Activity/Intervention |
|------|---|---|---|
| | <p>Grade 3 classes in public school (P.S) 86 and P.S. 148 that participated in LTA were randomly assigned to one of two treatment groups:</p> <p>1. Groups “tested” after they experienced the 20-week LTA program—to gauge the program’s effects on students (Treatment Group A).</p> <p>2. Groups “tested” after they experienced the 20-week LTA program and their teachers participated in LTA professional development—to gauge the effects of the professional development on students and teachers (Treatment Group B).</p> <p>3. As noted earlier, in 2004-05 third-grade classes in P.S. 94 and P.S. 149 served as the Control Group;</p> <p>4. 2005-06 P.S. 94 acted as the Control Group.</p> <p>What learning outcomes does the study measure/report?</p> <p>1. Academic achievement/skills:</p> <p>a) reading and discussing texts</p> <p>b) attitudes about class participation</p> <p>c) English Language Arts (ELA) test performance levels</p> <p>2. Cultural knowledge</p> <p>attitudes to art perceptions of artistic</p> | <p>3. 200 from the Control Group (no LTA program or professional development).</p> <p>Age of participants 6-10 years (52% 9 year-olds)</p> <p>Type of educational institution attended Secondary school 4 New York City public schools</p> <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants Students in the treatment and control groups were matched</p> <p>1. Students had similar demographic and socioeconomic characteristics.</p> <p>2. All did not meet the state performance standards on the English Language Arts test in 2001-02</p> <p>3. None of the students who were part of the study in these schools had participated in any Guggenheim Museum programs</p> <p>4. Three-quarters of students speak</p> | <p>classes “tested” who had not experienced the LTA program and whose teachers had not participated in LTA teacher professional development</p> <p>The comparisons were between 3 groups: Treatment Group A (LTA program only) vs Treatment group B (LTA program and teacher professional development) vs Control group (no LTA program or professional development)</p> <p>What is/are the setting(s) of the activity/intervention? Museum (art) Guggenheim museum</p> <p>Duration of the activity/intervention 3 months (and 1 day) to 6 months 20 weeks</p> <p>Intensity of the activity/intervention Unclear</p> |

| Item | Study | Sample | Activity/Intervention |
|------|--|--|-----------------------|
| | <p><i>process attitudes about art museums</i> <i>attitudes to learning Through Arts program</i></p> <p>3. Motivation for learning <i>attitudes to school</i></p> <p>4. Questionnaire/survey instrument completed by student <i>student performance measured using standardized questionnaires</i></p> <p>5. Questionnaire/survey instrument completed by teacher <i>Teachers' impact and performance were measured using standardized questionnaires</i></p> <p>6. One-to-one interview <i>student performance measured using interviews</i> <i>Teachers' impact and performance were measured using observations, and in-depth interviews</i></p> <p>6. Curriculum/subject-based test (assessment/exam etc completed by student) <i>test scores on the Third Grade Citywide English Language Arts (ELA) Test.</i></p> | <p><i>English and one other language at home (75 %). The most commonly spoken foreign language was Spanish (74 %)</i></p> <p><i>5. two statistically significant differences between the demographics of the Treatment and Control Groups. More Treatment Group students speak English at home than do Control Group students. For students who speak a non-English language at home, Spanish is spoken more frequently by Treatment Group students than</i></p> | |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|--|
| <p>Stanley J (2008)</p> <p>Weight of evidence D:</p> <p>Overall weight of evidence</p> <p>Medium</p> | <p>What are the broad aims of the study? <i>To assess the 'progress in writing attainment which can be associated with the operation of the Manchester Museums and Galleries Pilot (MAGPIE) programme during the academic year 2006-2007'</i></p> <p>What was the design of the evaluation? (Maryland Scale 3) Unmatched comparison group study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills Literacy</p> <p>Which methods were used to collect the data? School/college records (e.g. attendance, examination results) SATS</p> | <p>In which county/countries was the study conducted? UK</p> <p>Number of participants <i>410 intervention</i> <i>378 control</i></p> <p>Age of participants 6-10 years</p> <p>Type of educational institution attended Primary schools <i>Years 3 and 4</i></p> <p>Sex of participants Not stated</p> | <p>Which sectors does the engagement relate to? Arts and Museums</p> <p>Does the activity/intervention have a formal name? <i>Manchester Museums and Galleries Pilot (MAGPIE)</i></p> <p>What treatment/intervention did the control/comparison group receive? <i>Did not receive MAGPIE intervention</i></p> <p>What is/are the setting(s) of the activity/intervention? Arts and Museums</p> <p>Duration of the activity/intervention <i>The whole intervention took place over the period of one academic year</i></p> <p>Intensity of the activity/intervention <i>Once or twice</i></p> |

| Item | Study | Sample | Activity/Intervention |
|---|--|--|--|
| <p>Watson S, et al. (2007)</p> <p>Weight of evidence D:</p> <p>Overall weight of evidence</p> <p>Low</p> | <p>What are the broad aims of the study? <i>To investigate the impact that museums in the East of England region have on the attainment of secondary-age pupils</i></p> <p>What was the design of the evaluation? (Maryland Scale 1) Single-group post-test study</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills <i>Analysis of assignment marks based on the teachers' judgement of their progression</i></p> <p>Which methods were used to collect the data? 1. Questionnaire/survey instrument completed by student <i>pupils completed a questionnaire about their learning experiences after their museum visit</i> 2. Others <i>Observation of school visits to museums</i> <i>Interviews with teachers and pupils.</i> <i>Assignment marks from school records</i></p> | <p>In which county/countries was the study conducted? UK: East of England</p> <p>Number of participants <i>1. Assignment marks for 762 pupils were provided by the schools</i> <i>2. 451 pupils completed a questionnaire about their learning experiences after their museum visit</i> <i>2. 6 museums and 9 schools provided data</i></p> <p>Age of participants <i>aged 11-18 years (Key Stage 3 and above)</i></p> <p>Type of educational institution attended secondary school</p> <p>Sex of participants Mixed sex</p> | <p>Which sectors does the engagement relate to? Museums</p> <p>Does the activity/intervention have a formal name? No</p> <p>What treatment/intervention did the control/comparison group receive? one group only</p> <p>What is/are the setting(s) of the activity/intervention? <i>a number of diverse museums and galleries from across the East of England region</i></p> <p>Museum s(art, science, historical) <i>Including The Fitzwilliam Museum, Imperial War Museum, Duxford, Gressenhall Farm and Workhouse, Norfolk, Peterborough Museum, Cambridgeshire</i> <i>Verulamium, St Albans, Hertfordshire</i> <i>Suffolk Archives, Suffolk</i></p> |

| Item | Study | Sample | Activity/Intervention |
|------|--------------|--|---|
| | Case studies | <p>Please specify any other important features of the participants</p> <p>1. There was an over-representation of pupils from one school due to withdrawal at short notice of a selective grammar school, therefore 90% of the questionnaires were returned by pupils were from schools that visited Gressenhall Farm and Workhouse, and these schools provided a further 83% of the assignment marks that were returned</p> <p>2. Most schools (and pupils) were from rural areas and have below average % of pupils eligible for free school meals</p> | <p>Duration of the activity/intervention One visit to a museum or archive</p> <p>Intensity of the activity/intervention One visit</p> |

| Item | Study | Sample | Activity/Intervention |
|---|---|--|---|
| <p>West H (2008)</p> <p>Weight of evidence D:</p> <p>Overall weight of evidence</p> <p>Low</p> | <p>What are the broad aims of the study? <i>To investigate use of the TASC Wheel (Thinking Actively in a Social Context) in developing children's thinking skills and its' relationship to motivating learning (Abstract)</i></p> <p>What was the design of the evaluation? (Maryland Scale 2) one-group pre-post</p> <p>What learning outcomes does the study measure/report? Academic achievement/skills</p> <p>Which methods were used to collect the data? 1. Questionnaire/survey instrument completed by teacher 2. Self-completion report or diary <i>parents' diaries</i> 3. Observation 4. School/college records (e.g. attendance, examination results) SATs 5. Other <i>samples of children's writing</i> 6. Unclear</p> | <p>In which county/countries was the study conducted? UK</p> <p>Number of participants Unclear <i>3 schools were involved</i></p> <p>Age of participants 6-15 years (<i>years 2-6</i>)</p> <p>Type of educational institution attended Primary school</p> <p>Sex of participants Mixed sex</p> <p>Please specify any other important features of the participants <i>The research included children across a range of abilities</i></p> | <p>Which sectors does the engagement relate to? Museums</p> <p>Does the activity/intervention have a formal name? No</p> <p>What is/are the setting(s) of the activity/intervention? Museum (historical) <i>years 3 and 4 (school B) (STEAM museum)</i> <i>year 4 (school C) (Reading museum)</i></p> <p>Duration of the activity/intervention One day or less</p> <p>Intensity of the activity/intervention Once</p> |

| Item | Study | Sample | Activity/Intervention |
|------|--|--------|-----------------------|
| | <i>discussions with teachers and children (no further details)</i> | | |

General Appendices

General Appendix G1: The standard EPPI-Centre systematic review process

What is a systematic review?

A systematic review is a piece of research following standard methods and stages. A review seeks to bring together and 'pool' the findings of primary research to answer a particular review question, taking steps to reduce hidden bias and 'error' at all stages of the review. The review process is designed to ensure that the product is accountable, replicable, updateable and sustainable. The systematic review approach can be used to answer any kind of review question.

Stages and procedures in a standard EPPI-Centre review

Formulate review question and develop protocol.

Define studies to be included with inclusion criteria.

Search for studies (a systematic search strategy including multiple sources is used).

Screen identified potentially relevant studies against the inclusion criteria (inclusion criteria should be specified in the review protocol).

The results of screening (number of studies excluded under each criterion) should be reported.

Collect data in order to describe the characteristics of studies (keywording and/or in-depth data extraction).

Maintain bibliographic and review management data on individual studies.

Present descriptive information on each study in form of a 'map'.

At this stage the review question may be further focused and additional inclusion criteria applied to select studies for an 'in-depth' review

In-depth data extraction, including the results or findings of each study.

Assess study quality and relevance (the criteria used to make such judgements should be transparent and systematically applied).

Synthesise findings (the results of individual studies are brought together to answer the review question/s).

A variety of approaches can be used to synthesise the results. The approach used should be appropriate to the review question and studies in the review.

The review team interpret the findings and draw conclusions and implications from them.

Quality assurance

Quality assurance (QA) can check the execution of the methods of the review, just as in primary research; for example, through:

Internal QA: individual reviewer competence; moderation; double coding.

External QA: audit/editorial process; moderation; double coding.

Peer referee of: protocol; draft report; published report feedback.

Editorial function for report: by review specialist; peer review; non-peer review.

General Appendix G2: Inclusion and exclusion criteria**Stage 1: database creation**

| No. | Exclusion criteria | Inclusion criteria |
|------------|---|--|
| 1. | <i>Study not published in English.</i> | <i>Study published in English.</i> |
| 2. | <i>Study published prior to 1997.</i> | <i>Study published during or after 1997.</i> |
| 3. | <i>Study is off-topic (i.e., is not about people's engagement in culture and/or sport).</i> | <i>Study is about engagement in culture and/or sport.</i> |
| 4. | <i>Is not empirical primary research (e.g., opinion pieces, book reviews, bibliographies, news paper articles, editorials, strategy documents).</i> | <i>Is empirical primary research.</i> |
| 5. | <i>Is solely methodological research.</i> | <i>Study is not solely methodological research (i.e., include if study involves a methodological aspect (e.g., the aim is to validate an instrument), but findings about people's engagement in culture and/or sport are also reported).</i> |
| 6. | <i>Study is about engagement in cultural and/or sporting activities/sites, but not museums/libraries/archives or not as defined in the Taking Part survey (excludes, for example, studies about physical activity, walking, reading stories or poetry (as opposed to writing them), or architecture).</i> | <i>Study is about engagement in museums/libraries/archives, or cultural and/or sporting activities/sites as defined in the Taking Part survey.</i> |
| 7. | <i>Study investigates a sports or exercise-based rehabilitation/treatment programme for people with pre-existing, non-chronic, physical health problems (for example, post-surgical interventions).</i> | <i>Study investigates a sports or exercise-based rehabilitation/treatment programme for people with pre-existing mental health problem or with chronic health problems, such as arthritis or back pain.</i> |
| 8. | <i>Study only measures bio-medical outcomes (this criterion only applies to studies about engagement in sport).</i> | <i>Study measures any other outcome, in addition to bio-medical outcomes (this criterion only applies to studies about engagement in sport).</i> |

Stage 2a: mapping exercise/high quality studies

| No. | Exclusion criteria* |
|------------|--|
| 1. | <i>Out of scope (as defined by the selection criteria outlined in Table 2.1)</i> |
| 2. | <i>Not an impact study</i> |

| | |
|----|--|
| 3. | <i>Non-systematic review</i> |
| 4. | <i>No control group</i> |
| 5. | <i>Not about participation (in sporting activities) or about attendance or participation (in cultural activities); i.e., exclude studies about the remaining two engagement modes – deciding and producing – or studies about attendance at sporting events (either virtual or actual)</i> |
| 6. | <i>Not quantitative</i> |
| 7. | <i>Not pre-test/post-test</i> |
| 8. | <i>Non-relevant outcomes (e.g., health-related)</i> |

Stage 2b: mapping exercise/lower quality studies

| No. | Exclusion criteria* |
|-----|---|
| 1. | <i>Out of scope (as defined by the selection criteria outlined in Table 2.1)</i> |
| 2. | <i>Not impact study</i> |
| 3. | <i>Not about museums, libraries, archives or heritage</i> |
| 4. | <i>Non-systematic review</i> |
| 5. | <i>Not about attendance or participation in museums, libraries, archives or heritage activities/sites; i.e., exclude studies about the remaining two engagement modes – deciding and producing.</i> |
| 6. | <i>Not quantitative</i> |

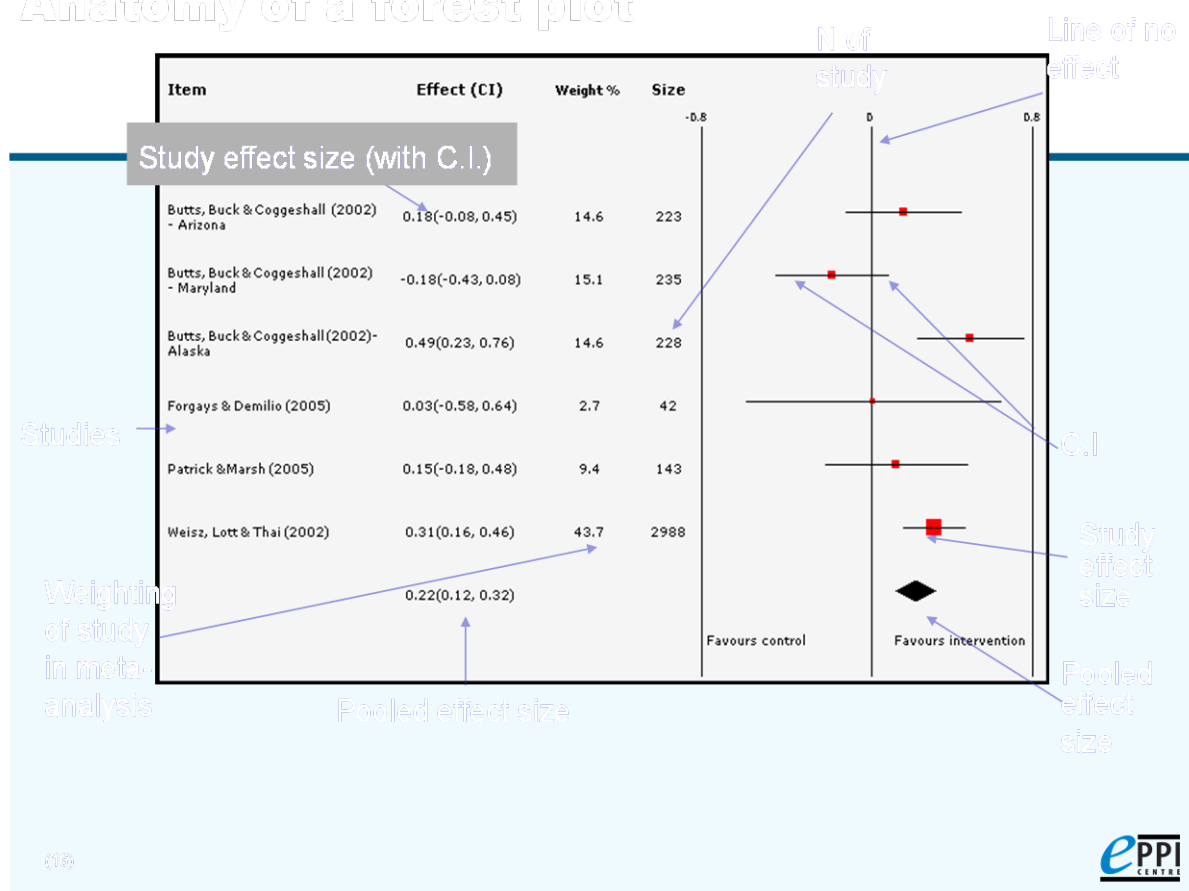
Stage 3: in-depth review (high quality studies)

| No. | Exclusion criteria |
|-----|--|
| 1. | Study did not focus on children/young people |
| 2. | Study did not measure at least one learning outcome |
| 3. | Learning outcome measured was related to the cultural/sporting activity (e.g., sports performance, musical skills, etc.) |
| 4. | Learning outcome measured was health-related knowledge/awareness (e.g., knowledge about STDs, HIV, etc.) |
| 5. | Not participation in sports activities |

* These studies were included in the in-depth review of research on learning outcomes for young people participating in the arts (Tripney et al., 2010b)

General Appendix G3: Forest plots

Anatomy of a forest plot



General Appendix G4: Study quality assessment framework

The following table outlines the procedure for calculating an overall quality rating (WoE D): for a fuller explanation, see the review technical report (Tripney et al., 2010a).

| WoE A: Quality of the execution of the study | WoE B: Study design | WoE C: Study relevance | WoE D: Overall quality rating |
|--|---|-------------------------------|--------------------------------------|
| <p>Was the knowledge of allocation to groups adequately prevented?</p> <p>Was incomplete outcome data addressed?</p> <p>Were the groups treated equally?</p> <p>Are the outcome measures reliable?</p> <p>Answers to individual questions were scored as follows and a total score calculated:</p> <p>yes = 1</p> <p>no = 2</p> <p>not relevant = 2</p> <p>unclear = 2</p> | <p>What is the design of the study?</p> <p>Answers were scored as follows:</p> <p>Randomised controlled trial (MSMS level 5) = score of 5</p> <p>Well-matched²⁰ comparison group pre-post test design (MSMS level 4) = score of 4</p> <p>Unmatched comparison group pre-post test design (MSMS level 3) = score of 3</p> <p>Single-group pre-post test design OR comparison group post-test only design (MSMS level 2) = score of 2</p> <p>Single group post-test only design (MSMS level 1) = score of 1</p> <p>If studies scored 5 or 4, then two additional questions were applied.</p> <p>Was the allocation sequence adequately generated?</p> <p>Was the allocation adequately concealed?</p> <p>If the answer was 'no' to either of these questions, 1 was subtracted from the</p> | | |

²⁰ Post-hoc statistical analysis used to control for differences between groups and comparison group considered to be well-matched to the intervention group on theoretically relevant factors (e.g., age, gender, etc).

| | score for that study. | | |
|---|---|--------------------|---|
| High = score of 4 Medium/High = score of 5 Medium = score of 6 Low/Medium = score of 7 Low = score of 8 | High = score of 5 Medium/High = score of 4 Medium = score of 3 Low/Medium = score of 2 Low = score of 1 | Fixed at 'High' | Not higher than WoE A or WoE B ²¹ |

²¹ A pre-established formula was used for moving from A, B and C to D. In this review, only A and B were taken into consideration and D could not be not greater than A or B (e.g., if A was medium/high and B was medium, then D would be medium).

General Appendix G5: Synthesis interpretation framework

| | |
|--------------------|--|
| What works | <p>These culture and/or sport programmes demonstrate evidence of beneficial impacts on participants compared to an alternative</p> <p>At least two evaluations of medium or greater quality with a positive pooled estimate of effect that excludes a result of no difference</p> |
| What does not work | <p>These culture and/or sport programmes demonstrate evidence of negative impacts on participants compared to an alternative.</p> <p>At least two evaluations of medium or greater quality with a negative pooled estimate of effect that excludes a result of no difference</p> |
| What is promising | <p>These are programmes where the level of certainty from available evidence is too low to support generalisable conclusions, but where there is some empirical basis for predicting that further research could support such conclusions</p> <p>Programmes are coded as promising if the pooled estimate of effect was found to be positive but did not exclude zero and there was at least one medium or better quality evaluation and the preponderance of the remaining evidence was also positive</p> |
| What is unknown | <p>Any programme not classified in one of the three above categories is defined as having unknown effects</p> |