Education Endowment Foundation

Chess in Schools

Evaluation report and executive summary

July 2016

Independent evaluators:

John Jerrim, Lindsey Macmillan, John Micklewright, Mary Sawtell and Meg Wiggins

Institute of Education

UCL

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

Danielle Mason
Head of Research and Publications
Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP

p: 020 7802 1679
e: danielle.mason@eefoundation.org.uk
w: www.educationendowmentfoundation.org.uk

About the evaluator

The project was independently evaluated by a team led by University College London, including John Jerrim, Lindsey Macmillan, John Micklewright, Mary Sawtell and Meg Wiggins.

The lead evaluator was John Jerrim.

e. j.jerrim@ucl.ac.uk

Contents

Executive Summary

The project

Chess in Primary Schools is a whole-school approach to teaching primary school children how to play chess. Children take 30 hours of chess lessons delivered by a tutor who is an experienced chess player, and the school is given the option to set up a chess club as a lunchtime or after-school activity. Chess classes are delivered during the school day and are expected to replace subjects such as music or PE.

The intervention was evaluated using a two-armed randomised controlled trial. The trial took place over the 2013/2014 academic year and assessed the impact of one year of Chess in Primary Schools on the mathematics attainment of pupils in Year 5. It was an effectiveness trial, with the intervention tested under realistic conditions in a large number of schools. This study looks at whether the intervention had an impact on attainment one year after the intervention had ended in June 2015. One hundred schools across 11 local education authorities (LEAs) in England participated in the trial, a total of 4,009 pupils. A process evaluation was also carried out to answer questions about implementation and to help explain the findings of the trial. The programme was delivered by the education charity Chess in Schools and Communities (CSC).

Key Conclusions
1. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
2. There is no evidence that the intervention had a positive impact on Key Stage 2 scores for children eligible for free school meals (FSM).
3. Although a current school teacher is allocated to every chess class, it is desirable for the tutors themselves to have strong class management and teaching skills. Without these, it was difficult to ensure that all children were suitably engaged in the chess lessons.
4. For successful implementation, class teachers need to work closely with the tutor and actively contribute to the intervention. It was felt that classes were less effective if the teacher did not actively take part, or was present only at the beginning and end of the class.
5. Half of the pupils who participated in the trial said that they liked the chess lessons a lot, and only 8% reported that they didn't like them. School teachers were very positive about the intervention and its impact on pupils' skills and behaviour.

How secure are the findings?

Security rating awarded as part of the EEF peer review process

Findings from this study have high security. The study was a large and well-designed clustered randomised controlled trial (RCT). It was an effectiveness trial, which means it aimed to test the intervention under realistic conditions in a large number of schools. Relatively few pupils were lost to the analysis and the pupils who were allocated to receive the intervention were similar to the pupils in the comparison group. There were no substantial threats to the validity of the results.

What are the findings?

Pupils, headteachers and class teachers were generally very positive about the intervention. In particular, pupils liked playing games of chess with their friends, and class teachers welcomed the enthusiasm of the tutors for sharing their expertise. School staff perceived that the chess lessons had a positive impact on maths ability, as well as on a range of important skills for learning such as concentration. What pupils liked least was tutors 'talking too much' and some teachers had concerns about the level of tutors' teaching skills. There were some departures from the intended delivery of the programme—primarily the level of class teacher engagement, which was lower than expected. Moreover, some schools reduced the number of maths lessons in the timetable in order to accommodate the chess lessons. Two key areas for intervention improvement emerged from the study. These were: (a) improving the teaching skills of tutors to help them keep all children engaged—specifically, improving their ability to manage difficult behaviour and manage classes where pupils had varying levels of ability; and (b) increasing the amount of tutor/class teacher liaison.

Despite the generally positive feedback received from schools from the process evaluation, the impact evaluation results found no evidence that the Chess in Primary Schools programme raised children's attainment in their Key Stage 2 exams. Indeed, the difference between the treatment and control arms was essentially zero. A similar impact was found for pupils eligible for free school meals, and for boys and girls. This is in contrast to the only other large-scale RCT of the impact of chess on educational attainment by Boruch and Romano (2011), who detected a substantial effect for primary school children in Italy and to another recent study by Gumede and Rosholm (2015), which found a positive effect of chess on primary school children's achievement in Denmark (effect size 0.15). The reasons for the differences could include the fact that this study measured the impact after one year, that this study used high stakes national tests, and the English setting.

How much does it cost?

The cost of delivering the intervention to two classes of Year 5 pupils is approximately £1,900, or £32 per pupil. The majority of this is to contribute towards CSC's costs of delivering the chess lessons (£1,200) and setting up the after-school chess club (£600).

Summary Table

Group	Effect size (95% confidence interval)	Estimated months' progress	Security rating	Cost rating
Pupils in CSC schools	0.01 (-0.15, 0.16)	0 months	🔒🔒🔒🔒🔒	**£**
Pupils in CSC schools eligible for FSM	0.01 (-0.18, 0.19)	0 months	🔒🔒🔒🔒🔒	**£**

Introduction

Intervention: The Chess in Primary Schools programme

The programme involved the CSC charity introducing chess lessons as part of a standard school day to Year 5 children within primary schools in England. The rationale behind the evaluation was that chess may help increase children's concentration, their ability to think strategically, and their self-confidence (see the 'logic model' in Figure 4 for further details). This would, in turn, lead to a long-term improvement in their academic achievement.

The intervention was delivered by fully trained tutors following a standardised 30-hour curriculum, details of which can be found at http://www.chessinschools.co.uk/sample_curriculum.htm. All chess tutors were chess specialists and did not necessarily have a teaching background. CSC regularly run one-day training courses aimed at anyone involved in school chess, including CSC tutors and class teachers. During the study intervention year CSC also ran a number of weekend seminars for tutors, at which the intervention was examined and a discussion process initiated on methods of enhancing the classroom delivery. Tutors were also able to exchange teaching methods and received useful tips on classroom management from some schoolteachers who attended.

Each participating school was asked to designate a teacher (or teaching assistant) who would help the CSC tutor to run the intervention in the class. This teacher or teaching assistant was required to attend a training seminar run by the CSC charity, and had full access to the CSC curriculum. The tutor was encouraged to discuss each lesson in advance with the class teacher / teaching assistant, in person or by email. Each school was also sent chess sets for classroom use, workbooks and curriculum books and also, later in the year, each child received a chess set and chess book.

The chess lessons were delivered as part of a regular school day. This meant that schools were to replace one regularly scheduled lesson to make room for the Chess in Primary Schools intervention. Schools were asked by CSC not to replace a maths or English lesson. Common lessons to be replaced were 'topic' or humanities, music and PE.

Whole-class teaching was used to deliver the Chess in Primary Schools programme. During lessons, material was presented using either a chess demonstration board or via the whiteboard. In order to use the whiteboard, each tutor was given specialist chess software, with the curriculum converted into a proprietary file format. Tutors had learning plans and objectives for each lesson, as well as worksheets for pupils. In each lesson, children shared a chess set on the desk to practise moves or, later, to play complete games in pairs. Tutors were encouraged to talk for no more than 15 minutes before allowing children to practise what they had been taught. In each school a chess club could optionally be set up at lunchtime or after school during the intervention period. Time at this chess club was additional to the 30 hours' taught curriculum time. Schools were encouraged to do this themselves. However, because of a lack of expertise, input from the CSC tutor was often required.

The game was taught piece by piece and visualisation of moves was required from lesson 2. By lesson 10, more abstract concepts such as 'check' and 'checkmate' were introduced. By the end of the first term, children were expected to be able to begin to play chess. Then, by the end of the second term, most children were expected to be able to play a reasonable game of chess. At the end of the school year, CSC organised competitions locally for groups of schools or within individual schools.

A 'business as usual' approach was used in control schools. This meant that no formal chess lessons were to be delivered (though if an after-school chess club already existed, this would continue to run). These schools were not allowed to access the intervention in either the 2013/14 or 2014/15 academic year. There was a small amount of crossover between treatment and control groups; six treatment schools did not deliver the intervention, while one control school gained partial access to the chess

treatment. This was accounted for with a contamination-adjusted intention to treat (CA-ITT) analysis to supplement the main intention to treat (ITT) analysis.

Box 1 provides a summary of the intervention, including details of the materials and procedures used, how the intervention was delivered, and the amount of chess instruction it was intended children would receive.

Box 1: TIDieR checklist

1. **Brief name**. Chess in Primary Schools

2. **Why: Rationale, theory and/or goal of essential elements of the intervention.** Chess would help to increase children's level of concentration, self-confidence and ability to think strategically. This would, in turn, lead to an improvement in their academic achievement.

3. **Who: Recipients of the intervention**. Year 5 (age 9/10) pupils within selected local education authorities in England.

4. **What: Physical or informational materials used in the intervention**. Chessboards, chess workbooks, chess software for whiteboard, CSC developed curriculum hanging demonstration board, classroom tables/chairs.

5. **What: Procedures, activities and/or processes used in the intervention**. Pupils are taught chess, as part of the school curriculum, by a trained chess tutor using the CSC curriculum. The CSC curriculum contains detailed 1-hour lesson plans that include mini-games and worksheets.

6. **Who: Intervention providers/implementers**. The intervention was provided by the charity Chess in Schools and Communities.

7. **How: Mode of delivery**. Face-to-face whole class delivery to children.

8. **Where: Location of the intervention**. Within primary school classrooms in England.

9. **When and how much: Duration and dosage of the intervention**. During the 2013/14 academic year. Children were to receive 30 chess lessons of 1 hour spread over the academic year.

10. **Tailoring: Adaptation of the intervention**. The tutors were provided with the CSC curriculum as the foundation for lessons, but were allowed to adapt lesson plans to suit individual classes.

Background evidence

The majority of studies that link chess to academic attainment have been conducted outside of England and include self-selecting intervention groups (e.g. Achiego et al., 2012). They have also tended to use 'low-stakes' tests (for which children in the control group are likely to be less motivated than those in the treatment group, as neither they nor their schools have anything riding upon the results). To our knowledge, only one randomised controlled trial of chess has been conducted (Boruch and Romano, 2011). This tested how 30 hours of chess tuition, provided by qualified tutors, influenced

8–9-year-olds' educational achievement in Italy. The study included 123 classes, randomly assigned to receive chess in either the 3rd or 4th grade. The intervention was found to increase mathematics achievement by an effect size of 0.34, though again 'low-stakes' tests were used. However, another recent quasi-experimental study by Gumede and Rosholm (2015) also found a positive effect of chess on primary school children's achievement in Denmark (effect size 0.15).

The rationale of this evaluation was to test the Chess in Primary Schools programme within the English setting. It was an effectiveness trial, with the intervention delivered at scale. As the Chess in Primary Schools programme is already well developed and widely used in schools, it was decided that a large-scale effectiveness trial was appropriate. The trial has a number of advantages over existing studies, including the use of 'high-stakes' tests, and a focus upon *medium term* effects of the intervention.

Evaluation objectives

The main question that the impact evaluation attempted to address was '*What is the impact of chess in schools on children's achievement in mathematics?*'. This was supplemented by a series of additional questions, including:

- What is the impact of teaching Year 5 children chess upon their Key Stage 2 reading and science test scores?
- What is the impact of teaching Year 5 children chess upon different mathematics sub-domains (e.g. mental arithmetic)?
- What impact does teaching FSM children how to play chess have upon their Key Stage 2 attainment?
- Does the impact of teaching chess differ between boys and girls?
- Is there any evidence that the Chess in Primary Schools programme has differential effects across the achievement distribution?

The process evaluation sought to answer the following questions:

- How feasible and acceptable is it for chess tutors to implement a 30-week classroom chess intervention in Year 5 of primary school? Could teachers who attended training and helped with the intervention continue to teach chess afterwards?
- How feasible and acceptable do teachers and headteachers feel it is for primary school children to play chess in class as part of the curriculum?
- What are the views, on the intervention, of the children who were offered it? How do these views vary by subgroup (e.g. boys vs. girls)?
- What are staff perceptions of the current and possibly sustained impact of the intervention on children's educational attainment? How do they think it affects different subgroups? How do they think it impacts on other matters such as class cohesion and school ethos? What are their perceptions of facilitators and barriers to impact? How scalable do they think the intervention is? What are their suggestions for change if the intervention was to be more widely implemented?

Project team

John Jerrim: Principle Investigator. Led the trial design, data analysis and writing of the final report. Overall management of the project.

Lindsey Macmillan: Assisted with trial design, data analysis and production of final report.

John Micklewright: Assisted with trial design.

Mary Sawtell: Joint-lead on the process evaluation design and analysis.

Meg Wiggins: Joint-lead on the process evaluation design and analysis.

Ethical review

The evaluation of the Chess in Primary Schools project was submitted to the Institute of Education ethics committee. Ethical approval was granted on 17 May 2013 (code FPS 504). School level consent has been obtained to conduct the trial and to access pupils' data from the National Pupil Database (NPD).

Trial registration

The protocol for this study is published online at:

http://educationendowmentfoundation.org.uk/library/chess-in-schools-protocol/

The trial has been registered with the independent ISRCTN website at:

http://controlled-trials.com/ISRCTN33648117

The trial registration number is ISRCTN33648117 and the DOI is 10.1186/ISRCTN33648117

Methods

Trial design

The study was designed as a clustered randomised controlled trial (RCT). At the start of the project, the evaluation team considered three options for randomisation of the intervention: (a) at the *pupil* level, (b) at the *class* level, and (c) at the *school* level. Option (a) was immediately ruled out due to the Chess in Primary Schools programme being designed as a group activity. We therefore focus on options (b) and (c).

Randomisation at the class level was deemed likely to be a powerful statistical design. (This was the approach taken in the Italian study of the impact of chess on attainment by Boruch and Romano, 2011.) However, the evaluation team decided this was outweighed by the following limitations.

First, concerns remained over possible 'contamination' between treatment and control classes. As each school would contain children in the same year in the two groups, it was deemed possible that children learning to play chess could encourage friends or siblings in the control group to also play chess outside of school. If chess does indeed have a positive effect on the outcome, such contamination would downwardly bias the estimated impact of the intervention. Second, parents may object to children receiving different 'types' of education within the same school as a result of random assignment of classes. For example, a parent who believes that chess will have a positive impact upon attainment may have been upset that their child had been assigned to the control group, when the child next door was getting the treatment in another class. Third, the need to alter the curriculum for one class but not another within the same year could present schools with an organisational problem. These second and third issues might have reduced the willingness of schools to take part in the trial, threatening both attrition and external validity.

Thus option (c) was chosen: a clustered randomised controlled trial, with randomisation at the school level. All forms within the selected year in a treatment school would receive the intervention; none would in the control schools. Moreover, control schools would continue to use 'business as usual' teaching, with the Chess in Primary Schools programme becoming available to them two years after the intervention began. This design is likely to provide less statistical power—but all three potential problems with class randomisation were likely to be greatly diminished.

Outcome measures

The primary outcome is pupils' Key Stage 2 maths test scores. KS2 scores are derived from a national examination that children sit in England at the end of primary school (when pupils are typically age 10 or 11). It is a reliable, externally valid measure that is a strong predictor of children's later educational outcomes. It is also a 'high stakes' test for schools, who are ranked in publicly available league tables by their pupils' performance. This test is not specific to the Chess in Primary Schools intervention and is marked blind to treatment. This outcome was pre-specified as part of the evaluation protocol. Maths was chosen because this is the academic area where Boruch and Romano (2011) reported a substantial effect.

Secondary outcomes include (i) performance in Key Stage 2 English tests, (ii) performance in Key Stage 2 Science tests (where available), and (iii) performance in sub-domains of the Key Stage 2 Maths test. The latter are known as 'paper A', 'paper B', and 'mental arithmetic', with the following links providing the three test papers that children took in June 2015:

Paper A:
http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_paper1_PDFA.pdf

Paper B:
http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_paper2_PDFA.pdf

Mental arithmetic:
http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_mentalmathematics_transcript_PDFA.pdf

Note that the Chess in Primary Schools intervention was delivered while children were in Year 5 (age 9/10). Key Stage 2 tests (the outcome) were conducted at the end of Year 6. Hence outcomes have been measured one year after the intervention finished. The trial has therefore been designed to detect a *medium term* effect of the intervention.

Baseline test

Children's Key Stage 1 (KS1) maths, reading, writing and science test scores were used to measure children's academic achievement prior to the Chess in Primary Schools intervention. These are based upon teacher assessments of pupils when they were age 7—and thus before schools were randomly assigned to treatment and control groups. Indeed, at the point these baseline tests were conducted, teachers would have been unaware that the Chess in Primary Schools trial would take place. These baseline scores are used to (i) investigate balance between treatment and control groups in terms of prior attainment, and (ii) increase power and reduce any imbalance between treatment and control groups in the statistical analysis.

Participant selection

The Institute of Education (IoE) and Chess in Schools and Communities (CSC) teams first identified specific local education authorities (LEAs) in England where CSC had capacity to deliver the intervention. The LEAs selected were:

- City of Bristol
- Hackney
- Hammersmith and Fulham
- Leeds
- Liverpool
- Middlesbrough
- Newham
- Sefton
- Sheffield
- Southwark
- Tameside

The Institute of Education then produced a list of all primary schools within these LEAs. Private schools and schools where CSC already operated were excluded. For logistical reasons, it was also agreed that any primary school with four-form entry would be excluded from the evaluation. Schools with more than 90 pupils aged 11 were thus removed from the sampling frame. This was working on the assumption that there were approximately 30 pupils

per class within primary schools, and that year group size within schools would not significantly change within a short space of time.

The sampling frame was further restricted to schools with a high intake of disadvantaged pupils, based upon the percentage of children receiving free school meals (FSM). Schools were only selected if at least 37% of their children had either been eligible for FSM within the last six years or had been looked after by the local authority continuously for six months[1]. Thus the population of interest was defined as Year 5 state school pupils within the selected LEAs, who attended a one, two or three form entry primary school, which had a high proportion of disadvantaged pupils, and was not currently enrolled in the Chess in Primary Schools programme.

This final list contained 442 schools. CSC were then asked to recruit 100 of these schools by the third week of July 2013. CSC sent all schools an information pack. Those that agreed to take part in the trial completed a consent form to participate in the study and to allow access to data from the National Pupil Database (NPD). Ninety-two schools were recruited into the trial by this date. A further eight schools were recruited by September 2013, and were also included in the trial (bringing the total to 100). School-level consent to participate in the trial, and to allow the evaluation team access to the NPD data, was obtained from schools prior to randomisation. All children in the Year 5 treatment schools were required to participate in the programme to avoid selection problems.

Sample size

The evaluation team regarded 100 schools as the minimum necessary to detect an effect of approximately 0.18 of a standard deviation in Key Stage 2 (KS2) mathematics test scores. This calculation assumed:

i. an intra-cluster correlation (ICC) of 0.15 at the school level[2];
ii. equal cluster sizes of 60 Year 5 pupils per school[3];
iii. 40% of the variation in KS2 maths test scores would be explained by baseline covariates[4]; and
iv. 80%t power for a 95% confidence interval.

Table 1 provides estimates of the ICC for the actual sample of schools/pupils that took part in the study. Estimates are presented for baseline (KS1 average points score) and follow-up (KS2 maths) tests, when using either a fixed or random school-level effect. The ICC for KS1 average point scores (APS) was 0.08 when using a fixed effects model. The analogous ICC for KS2 maths was 0.13. In the results section, we illustrate that 45% of the variance in KS2 maths test scores can be explained by the baseline covariates. Using these figures in place of (i) and (iii) above, we calculate the minimum detectable effect in this trial was approximately 0.16 (see Table 3 below for further details).

[1] The figure of 37% was decided upon so that the population list given to CSC would contain a population of approximately 450 schools from which they could recruit into the trial.
[2] A value of 0.15 for the ICC was chosen after the team conducted an analysis of within and between school variation in Key Stage 2 test scores within the National Pupil Database.
[3] The figure of 60 pupils was based on the assumption of most recruited schools being two-form entry, with each form containing 30 pupils.
[4] A value of 0.4 was chosen after the team conducted an analysis of the association between Key Stage 1 and Key Stage 2 test scores within the National Pupil Database.

Table 1: Estimated inter-cluster correlation

	Fixed effect	Random effect
Key Stage 1 APS	0.08	0.05
Key Stage 2 Maths	0.13	0.11

Note: Figures refer to the proportion of the variation in pupils' test scores occurring between schools.

Randomisation

The trial was designed as a stratified, clustered randomised controlled trial—with random allocation occurring at the school level. Schools were first separated (stratified) into groups based upon (i) the percentage of pupils achieving level 4 or above at the end of KS2 in both English and mathematics, and (ii) the percentage of current KS2 pupils who had been eligible for FSM in the last six years. The schools were categorised into three strata for each variable and then the variables were cross-tabulated to create the following nine strata:

1. Low achieving–low FSM
2. Low achieving–middle FSM
3. Low achieving–high FSM
4. Middle achieving–low FSM
5. Middle achieving–middle FSM
6. Middle achieving–high FSM
7. High achieving–low FSM
8. High achieving–middle FSM
9. High achieving–high FSM

A tenth stratum was then included to incorporate the eight schools that were recruited into the trial between the end of July and September 2013:

10. 'Late' recruited schools

The number of schools within each stratum can be found in Table 2.

Table 2: Number of schools within each stratum

Strata ID	Average achievement	% Free school meals	Schools per strata
1	Low	Low	7
2	Low	Medium	12
3	Low	High	12
4	Medium	Low	13
5	Medium	Medium	11
6	Medium	High	7
7	High	Low	11
8	High	Medium	10
9	High	High	9
10	Late recruitment	Late recruitment	8

Within each stratum a random number was then drawn from a uniform distribution for each school. The schools in the bottom half of the random draw distribution, *within each stratum*, were assigned to the control group. Schools with a number in the top half of the random draw distribution were assigned to treatment. If the stratum contained an odd number of schools,

then the school with the median random draw was randomly assigned to treatment or control. STATA version 12 has been used to generate all random numbers. Note that all schools in strata 1 to 9 were randomly assigned on the same day in the third week of July 2013. Schools in stratum 10 were randomised on a separate day in August 2013.

The creation of the random number sequence and allocation of participants was done by Dr John Jerrim.

Analysis

The analysis strategy used intention to treat. Analysis of whether the intervention was effective or not was based upon the following OLS regression model:

$$Y_{ij}^{Post} = \alpha + \beta.Treat_j + \gamma.Y_{ij}^{Pre} + \varepsilon_{ij} \quad (1)$$

where:

$Y^{pre} =$ children's KS1 test scores in maths, reading, writing and science

$Y^{post} =$ children's KS2 maths test score

Treat = a binary variable indicating whether the child was enrolled in a treatment or control school (0 = control; 1 = treatment).

ε = error term (with children clustered within school)

i = child i

j = school j

To account for the clustering of pupils within schools, the STATA survey (svy) command is used to make Huber-White adjustments to the estimated standard errors. The coefficient of interest from equation 1 is $\boldsymbol{\beta}$ – is there a positive effect of the Chess in Primary Schools treatment?

After our main analysis, we re-estimate model 1 (i) separately for boys and girls, and (ii) separately for FSM pupils. The same analysis process has been followed for the secondary outcomes (sub-components of the Key Stage 2 maths tests, Key Stage 2 English scores, Key Stage 2 Science scores)[5].

The evaluation team has also conducted an 'on-treatment' analysis, where we investigate whether the effectiveness of the intervention varies by how it was implemented within schools. This part of the analysis was not pre-specified in the study protocol, but was undertaken in order to investigate whether there was any difference in the effect of the treatment by the fidelity of the treatment. In order to conduct this analysis, children within treatment schools were asked how much they liked the chess lessons that were delivered. Each chess tutor was then assigned to one of three categories (high, medium, low) depending upon the proportion of children that they taught who responded that they 'liked the lessons a lot'[6]. We label this variable 'chess tutor quality'. In our analysis, mean post-test

[5] If an observation is missing Key Stage 2 test score data, it has been excluded from the analysis.

[6] If less than a third responded positively, the tutor was assigned to the low group. If between one-third and two-thirds responded positively, the tutor was assigned to the middle group.

scores for children taught by teachers within these three 'tutor quality' groups are compared to mean post-test scores for the control group. The intuition is that liking chess is one of the key 'change mechanisms' through which we anticipate an effect to occur; thus greater levels of pupil enjoyment is likely to indicate a more fertile treatment. In an additional 'on-treatment' analysis, we also illustrate whether the intervention was more effective when the class teacher attended the CSC one-day training workshop (as anticipated in the study protocol).

In addition to the intention-to-treat (ITT) analysis described above, we also undertake a contamination adjusted intention-to-treat analysis. This was done in order to test the sensitivity of our results to the small amount of crossover between treatment and control groups that occurred during the trial. Further details are provided below.

Costs

Information of costs was gathered directly from the programme developer (i.e. the charity Chess in Schools and Communities). Specifically, the evaluation team asked CSC a series of questions regarding various aspects of the resources needed to run the intervention (see Appendix B for a list of questions asked).

The response of CSC to these questions was then used to calculate the cost of a school participating in the programme next academic year. This figure varies by the size of the school, as CSC requests a larger contribution to their overheads from bigger schools. As regular class teachers were expected to complete a one-day training course, a figure of £200 was added to this value for each teacher from a school who would attend (under the assumption this would cover the costs of employing a supply teacher)[7]. These figures were then added together to give a total cost. The total cost was then divided by the number of pupils, under the assumption of 30 pupils per school form.

Implementation and process evaluation

The process evaluation was integral to the trial. It was designed with three key purposes: to assess the fidelity of delivery of the intervention, to answer questions related to the feasibility of the intervention, and to support understanding of the results of the impact evaluation.

Constructing a logic model

A logic model was developed to clarify assumptions on CSC's views of the theory of change of the programme and to provide a framework to support the evaluation including the assessment of fidelity and explanation of findings.

The construction of a logic model was undertaken by the research team, using a three-stage Delphi consultation exercise, designed to achieve consensus within a group of eight experts in the CSC programme. The consultation was carried out by email without individuals conferring or seeing the responses of others in the group. The first stage asked for views on what the different components of the logic model were at each stage of the causal pathway. The components submitted by the participants were consolidated by the research team. The eight participants were then asked to rank the listed components in order of importance. The research team analysed the ranked lists and constructed a draft logic model that reflected the combined views of the experts. This draft was emailed to the participants with a request for

Teachers in the high group had at least two-thirds of children reporting that they liked the chess lessons a lot.

[7] Following EEF guidance, we spread this cost over three years.

any amendments to be provided. The final version of the logic model can be found in the process evaluation results section (see Figure 4).

Pre-intervention data collection

Baseline headteacher survey

A headteacher survey was conducted with all the schools enrolled in the study in July 2013, immediately prior to randomisation. This short, paper-based survey asked about any current or recent chess playing in the school, headteachers' own chess playing experience and their keenness for taking part in the trial. The rationale for this survey was to be able to assess the general level of chess interest and exposure in each school prior to intervention delivery. Non-respondents were sent two email reminders which included an invitation to complete the survey with a researcher over the telephone.

Observation of CSC training for tutors and teachers

During the set-up phase of the trial, two members of the research team observed, together, a CSC one-day training course, aimed at teachers, prospective tutors and anyone involved in school chess. The two researchers then carried out one observation each of two further one-day training courses to which teachers from treatment schools had been invited. Free-form observation notes were taken by the researchers.

Data collection during and immediately after intervention delivery

Observations of chess lessons

Observations were carried out in four schools approximately half-way through the intervention delivery period (March/April 2014). The four schools were purposefully selected to ensure a range of the following:

- location in the country;
- number of classes in a year group;
- previous levels of chess exposure and interest (as assessed by the baseline headteacher survey); and
- tutor factors (general teaching experience; current level of chess playing; number of years employed by CSC; whether they worked individually or as a pair; and gender).

The aims of the non-participatory observations were to provide information on: how the intervention was delivered, with a particular focus on fidelity; the acceptability by all stakeholders; and barriers and facilitators to delivery. Two evaluation team researchers carried out observations of one-hour chess lessons in eight Year 5 classes in the four selected schools. The researchers completed a semi-structured proforma during the observation, which included prompts for the various inputs listed in the logic model. Immediately after the lesson the observer had a brief discussion with the class teachers and the CSC tutor to clarify any issues arising from the observation.

Teacher survey

An online survey of all Year 5 teachers in intervention and control schools was carried out in June and July 2014 when intervention delivery was nearly complete. Teachers in schools in both trial arms were asked questions about themselves (e.g. gender and years of teaching experience); their class, including the amount of support they had in the classroom; other interventions during the year aimed at raising maths and literacy attainment; and numbers of

pupils with particular needs such as special educational needs (SEN) and English as an additional language (EAL). Intervention teachers were also asked questions on the acceptability, feasibility and sustainability of the intervention and on their perceptions of impact. Two email reminders were sent to all non-responders which offered the option of completion over the telephone. An additional paper version of the questionnaire was sent by post with an accompanying prepaid reply envelope to all remaining control teachers who had not responded.

Stakeholder interviews

In-depth telephone interviews were conducted during July 2014 with two tutors, two headteachers, and five class teachers (from four different treatment schools). These were audio taped with the permission of the interviewee. Participants were purposively selected, based on survey responses, to provide insights into key themes emerging from the survey and observation data. Notes were made during and immediately after the telephone interview to capture the key points. The audio recording was used as a check where there was uncertainty or to extract a particular quote. A further five teachers, two headteachers and five tutors were briefly interviewed, face to face, during observation site visits. Notes of key points were made by the researchers immediately after these discussions.

Observation of CSC seminars for tutors

A further evaluation activity was a one-day observation at each of two weekend seminars organised by CSC for tutors. The aim of these seminars was to support tutors in developing their teaching and classroom management skills. The programmes for these events included presentations by external experts, and tutors and others involved with CSC sharing personal experiences and tips. Free-form notes were taken by the researcher observing the seminars.

Data collection in the year following intervention delivery

Pupil survey

A short pupil survey was carried out in treatment schools in February 2015, approximately seven months after the chess lessons ended. Participation was optional for the students, and parents were given the opportunity to opt them out of this exercise. The survey was paper based and self-completion, with administration by class teachers. Packs of questionnaires, with accompanying guidance for teachers, were sent to the current teacher (Year 6) of each class that had received the chess lessons the previous year. The survey included closed questions on acceptability of the lessons; chess playing prior to the lessons; any chess playing since the chess lessons ended. Free text boxes were provided for pupils to write about their views on the best and worst aspects of the chess lessons.

Interview with intervention provider

A face-to-face semi-structured audio-taped interview with a senior member of the CSC head office team was conducted in May 2015. The main aim of this final data collection exercise was to explore themes that had emerged from other data sources from the perspective of the providers of the intervention. Questions asked covered: views on the process of overseeing the delivery of the intervention (including training and support for tutors); what was learned from the process; and whether any associated subsequent changes had been made or planned. As with the other stakeholder interviews, notes were made immediately after the interview. The tape recording was used for reference as required.

Analysis

Framework analysis was used for the analysis of the qualitative data from interviews and observations. This involved constructing frameworks based on key themes that answered the main research questions. This method allowed exploration of the data by both theme and respondent-type, enabling identification of patterns and associations across themes and types of respondents.

Descriptive statistical analyses of the teacher, headteacher, tutor and pupil surveys was carried out using SPSS V22. Chi-square tests were used to measure statistical significance.

Using data from across process evaluation sources, measures of intervention dose and quality were constructed for each school.

Timeline

Date	Activity
June 2011	Sample children's Key Stage 1 tests conducted
March – July 2013	Schools recruited (8 late schools recruited in August 2013)
July 2013	Schools assigned to treatment or control group (8 additional schools in August 2013).
October 2013 – July 2014	Chess in Primary Schools programme delivered in treatment schools
June 2015	Key Stage 2 (post-tests) conducted
October 2015	Analysis conducted

Impact evaluation

Participants

Sample allocation

Figure 1 provides details of sample allocation and attrition. One hundred schools were recruited to participate in the trial. Schools were randomly allocated to treatment (n = 50) and control (n = 50) groups.

All Year 5 children enrolled in the 100 participating schools in the trial on 3 October 2013 were considered to be part of the Chess in Primary Schools trial. (This was the date of the autumn school census in 2013.) Information on school enrolment on this date was drawn directly from the National Pupil Database (NPD). A total of 1,954 children were enrolled in the 50 control schools and 2,055 in the 50 treatment schools.

Missing data at baseline

Pupils' KS1 maths, reading, writing and science test scores were taken directly from the NPD. Information was missing for a small number of pupils who were not enrolled in a school in England at age 7 or where there were problems linking NPD data over time[8]. KS1 data was available for a total of 3,775 (94%) of the 4,009 children within the 100 participating schools. A 'missing' dummy variable is included in the OLS regression model to ensure these observations are not dropped for our analysis.

Attrition between intervention and post-test

The schools and children recruited into the trial were tracked using the NPD. Pupils who moved to a different school could be tracked via their unique pupil number (UPN) and were included in the final analysis. KS2 test score data could be linked for 3,865 of the 4,009 pupils initially recruited into the trial (see Figure 1). This group of pupils forms our final analysis sample.

Contamination

Six out of the 50 schools assigned to the treatment group dropped out of the Chess in Primary Schools programme before the intervention had begun.

One control school was unwilling to accept their assigned group and delivered chess lessons to their Year 3 pupils. Although chess lessons were not provided to the Year 5 pupils who were the intended controls, there is nevertheless an element of non-compliance.

To summarise, six of the schools who were meant to receive the Chess in Primary Schools treatment did not, while one control school managed to (partially) gain access to the intervention. As per our study protocol, our main analysis will follow an intention-to-treat (ITT) approach. This is where treatment and control groups are defined based upon their initial random allocation. However, we also present alternative estimates applying a contamination adjusted intention-to-treat (CA-ITT) methodology. This is an instrumental variable (IV) approach, where initial treatment/control allocation is used as an IV for actual receipt of the intervention. The key assumption is that initial random allocation (the IV) is strongly

[8] Note that children who move between schools can be tracked through the NPD—so children are not lost from the trial for this reason.

associated with the probability of actually receiving the intervention, but is not independently associated with the outcome (KS2 scores). This assumption is likely to hold as the extent of non-compliance is relatively small, meaning that initial allocation will strongly predict who actually received the treatment, and there is no reason to believe the IV and the outcome are associated (as the IV is random assignment to treatment/control status). The CA-ITT methodology also assumes that that if non-compliers had received the treatment, the treatment would have had the same effect as it did on the compliers.

It is important to recognise that ITT and CA-ITT address two different (though related) questions. Whereas ITT asks: *How much do study participants benefit from being assigned to a treatment group?*, CA-ITT considers: *What is the size of treatment benefit for someone who receives the treatment?* In other words, CA-ITT attempts to abstract from the problem of contamination. Thus a benefit of CA-ITT is that it leads to improved accuracy in estimating the size of treatment benefit for individuals who receive the treatment (Sussman and Hayward, 2010).

Figure 1: Participants in the Chess in Primary Schools trial

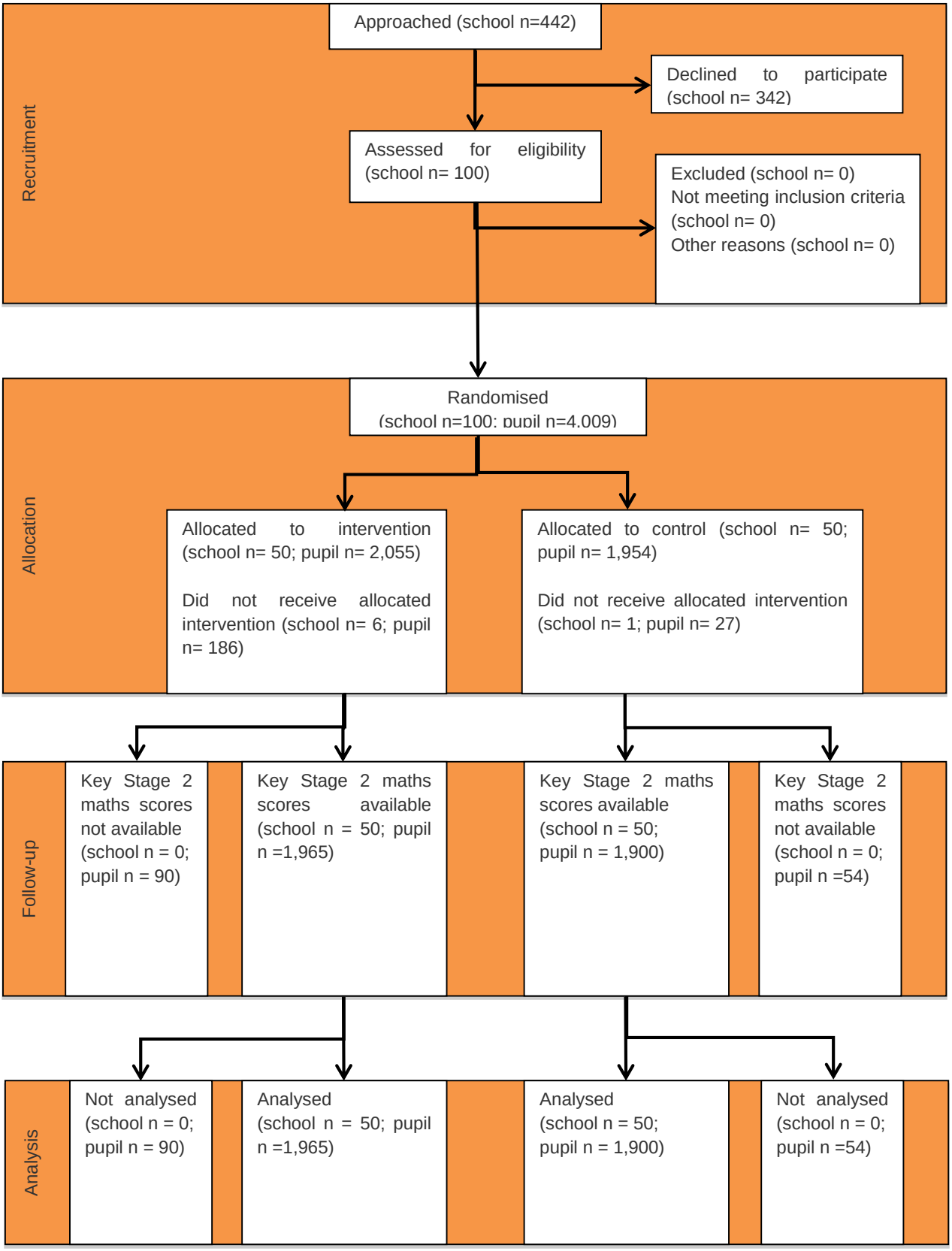


Table 3: Minimum detectable effect size at different stages

Stage	N [schools/pupils] (n=intervention; n=control)	Correlation between pre-test (+other covariates) & post-test	ICC	Blocking/stratification or pair matching	Power	Alpha	Minimum detectable effect size (MDES)
Protocol	100 schools (50 treatment, 50 control). 6,000 pupils (3,000 treatment,3,000 control)	0.63 (40% of variance explained)	0.15	10 Stratum based upon FSM and prior achievement	80%	0.05	0.18
Randomisation	100 schools (50 treatment, 50 control). 4,009 pupils (2,055 treatment,1,954 control)	0.67 (45% of variance explained)	0.11	10 Stratum based upon FSM and prior achievement	80%	0.05	0.16
Analysis (i.e. available pre- and post-test)	100 schools (50 treatment, 50 control). 3,865 pupils (1,965 treatment,1,900 control)	0.67 (45% of variance explained)	0.11	10 Stratum based upon FSM and prior achievement	80%	0.05	0.16

Note: Correlation between pre-test and post-test based upon OLS regression model including controls for gender, FSM, KS1 maths score, KS1 reading score, KS1 writing score, and KS1 science score.

Pupil characteristics

Table 4 compares KS1 scores for children in the treatment and control groups across four subject areas (numeracy, reading, writing, and science). All children for whom KS1 information could be linked are included in this comparison. The distribution of KS1 maths scores is very similar across the two groups, with differences at any given level typically just one or two percentage points. Similar findings hold for KS1 reading and writing. Indeed, the only instance where there is a difference of meaningful magnitude is KS1 science, where more children reach level 3 in the treatment group (14%) than in the control group (7%). We have additionally looked at mean KS1 average point scores (APS) for treatment and control groups. The difference is again small, standing at 0.05 standard deviations. Overall, Table 4 suggests that the sample is well balanced in terms of prior academic achievement.

Table 5 considers balance between treatment and control groups in terms of other observable characteristics. (These characteristics are presented for all children initially randomised.) There is broadly the same proportion of boys and girls in the two arms of the trial, though with slightly more children eligible for FSM in the control group (36%) than the treatment group (33%). Nevertheless, most of the differences observed between treatment and control groups in Table 5 are relatively small. Overall, Table 5 suggests that the treatment and control groups are also reasonably well balanced on a range of baseline characteristics.

Table 4: Comparison of baseline (Key Stage 1) test scores between treatment and control groups

Variable	Intervention group		Control group	
	n	Percentage (or standardised mean)	n	Percentage (or standardised mean
Key Stage 1 maths				
Level 1	242	12%	236	12%
Level 2A	441	21%	450	23%
Level 2B	590	29%	567	29%
Level 2C	366	18%	356	18%
Level 3	246	12%	191	10%
Missing	170	8%	154	8%
Key Stage 1 reading				
Level 1	330	16%	309	16%
Level 2A	428	21%	457	23%
Level 2B	523	25%	491	25%
Level 2C	278	14%	280	14%
Level 3	304	15%	243	12%
Missing	192	9%	174	9%
Key Stage 1 writing				
Level 1	363	18%	373	19%
Level 2A	340	17%	319	16%
Level 2B	586	29%	509	26%
Level 2C	433	21%	453	23%
Level 3	116	6%	112	6%
Missing	217	11%	188	10%
Key Stage 1 science				
Level 1	306	16%	297	16%
Level 2	1,317	68%	1,369	74%
Level 3	266	14%	131	7%
Missing	166	2%	157	3%
Key Stage 1 average point score				
Standardised mean	1,932	0.024	1,843	-0.025
Missing	123		111	
School n				
Pupil n				

Notes: All figures refer to percentages, except KS1 average points score (which has been standardised to have a mean of 0 and standard deviation of 1 across the participating sample of 4,009 pupils). All analysis performed at the pupil level (i.e. figures refer to percentage of pupils – not percentage of schools).

Table 5: Comparison of demographic characteristics between treatment and control groups

Variable	Intervention group		Control group	
	n	Percentage	n	Percentage
Eligible for FSM				
No	1,376	67%	1,250	64%
Yes	679	33%	704	36%
Gender				
Female	1,007	49%	997	51%
Male	1,048	51%	957	49%

Notes: All figures refer to percentages. All analysis performed at the pupil level (i.e. figures refer to percentage of pupils—not percentage of schools). There is no missing data for these variables.

External validity

Schools were not randomly selected into the trial. Rather, the evaluators composed a list of 442 schools within the 11 local authorities who were eligible to participate in the trial (see 'Method' section above). This list of schools was then given to the CSC project team, who were asked to recruit 100 schools to participate in the trial. Put another way, the CSC team had to ensure that at least 22%e of the 442 eligible schools were recruited.

Table 6 considers whether pupils within the 100 participating schools have similar baseline (KS1) test scores to pupils in the population of 442 schools who were eligible to take part in the trial. (Figures for all state school pupils in England are also provided for context, though the trial has not been designed to generalise to the country as a whole. This data has been drawn from the National Pupil Database.)

The percentage of children in each Key Stage 1 performance level is very similar across the 'trial participants' and 'eligible' samples. Standardised APS scores differ by less than 0.01 standard deviations between these two groups. A similar finding holds for the distribution of KS1 levels across each of the four subject areas; differences between trial participants and the eligible population is never more than one or two percentage points. Thus, despite the absence of random sampling, children who took part in the trial were very similar to the population of pupils they were meant to represent in terms of prior academic achievement.

Table 7 presents a similar comparison for other demographic characteristics. There are slightly fewer children with English as an additional language (EAL) among trial participants (34%) than in the eligible population. Likewise, London is somewhat over-represented compared to the rest of the country. However, differences observed between eligible and participating pupils are nevertheless relatively small in terms of magnitude. Overall, this reinforces the main message of Table 6—the sample of trial participants is broadly representative of the population who were eligible to take part (at least in terms of observable characteristics).

Table 6: Comparison of Key Stage 1 test scores of trial participants to (i) the population of eligible pupils, and (ii) all state school pupils in England

Variable	Trial participants	All eligible pupils	England
Key Stage 1 maths			
Level 1	12%	12%	8%
Level 2A	22%	24%	27%
Level 2B	29%	30%	27%
Level 2C	18%	20%	15%
Level 3	11%	11%	20%
Missing	8%	3%	2%
Key Stage 1 reading			
Level 1	16%	16%	12%
Level 2A	22%	23%	25%
Level 2B	25%	27%	23%
Level 2C	14%	14%	12%
Level 3	14%	15%	26%
Missing	9%	4%	3%
Key Stage 1 writing			
Level 1	18%	20%	15%
Level 2A	16%	16%	20%
Level 2B	27%	29%	29%
Level 2C	22%	23%	20%
Level 3	6%	7%	13%
Missing	10%	5%	4%
Key Stage 1 science			
Level 1	15%	16%	10%
Level 2	67%	72%	68%
Level 3	10%	10%	20%
Missing	8%	2%	2%
KS1 average points score (standardised across the population in England)	-0.28	-0.29	0.00
School n	100	442	0
Pupil n	4,009	16,397	570,344

Notes: 'All eligible pupils' refer to all pupils in the schools that were eligible to be recruited into the trial. Trial participants includes both treatment and control group. England provides figures for all state school pupils. In this table, KS1 average points score has been standardised across the 570,344 pupils in the English state school population. Hence, for this variable, figures will not match between Table 4 and Table 6.

Table 7: Comparison of demographic characteristics of trial participants to (i) the population of eligible pupils, and (ii) all state school pupils in England

Variable	Trial participants	All eligible pupils	England
Eligible for FSM			
No	66%	65%	82%
Yes	35%	35%	18%
Gender			
Female	50%	50%	49%
Male	50%	51%	51%
Language Group			
English	65%	63%	82%
Other	34%	37%	18%
Local Authority			
Hackney	15%	10%	-
Hammersmith and Fulham	6%	4%	-
Southwark	17%	11%	-
Newham	13%	14%	-
Liverpool			
Sefton	5%	4%	-
Tameside	7%	5%	-
Sheffield	4%	9%	-
Leeds	16%	18%	-
Bristol	6%	6%	-
Middlesbrough	4%	5%	-
Ethnic Group			
White	52%	54%	77%
Black	22%	19%	5%
Asian	12%	14%	10%
Mixed	8%	7%	5%
Other	4%	4%	2%
Unclassified	1%	1%	1%
Chinese	0%	1%	0%
School n	**100**	**442**	**0**
Pupil n	**4,009**	**16,397**	**571,733**

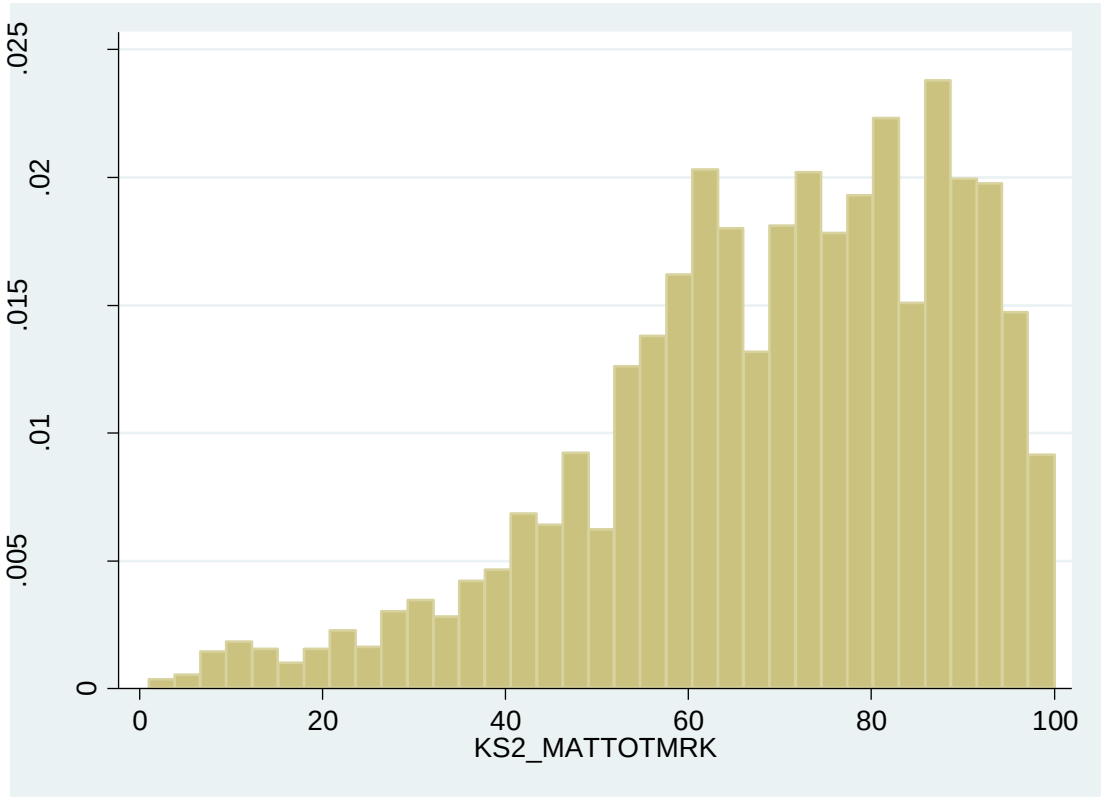
Notes: 'All eligible pupils' refer to all pupils in the schools that were eligible to be recruited into the trial. Trial participants includes both treatment and control group. England provides figures for all state school pupils.

Outcomes and analysis

Descriptive statistics

Figure 2 plots the distribution of Key Stage 2 test scores for the children in the analysis sample. There is little evidence of either floor or ceiling effects, though the distribution does have notable negative skew. The overall mean is 70 points, and the standard deviation is 20. We have also estimated the strength of the association between children's Key Stage 1 average points score and their marks in the Key Stage 2 maths exam. The correlation is 0.65, with around 40% of the variance in Key Stage 2 maths scores explained.

Figure 2: The distribution of children's Key Stage 2 raw scores



Notes: The y-axis refers to the probability density. The x-axis refers to the total score on the Key Stage 2 mathematics test.

Primary outcome: Overall Key Stage 2 maths scores

Results are presented in Table 8. The first row presents the intention-to-treat (ITT) estimates, while the second provides the contamination adjusted intention-to-treat (CA-ITT) estimates. Children who received the Chess in Primary Schools intervention achieved Key Stage 2 maths scores no higher than the control group, with an effect size of 0.01 and 95% confidence interval ranging from -0.15 to +0.16. Similar substantive conclusions hold for both the ITT and CA-ITT analyses. In additional analysis (results not presented), we have also re-estimated the effect of the intervention having excluded the seven schools that removed a maths lesson in order to make room for the CSC curriculum. The effect size actually fell slightly, to -0.02 (95% confidence interval from -0.18 to +0.13), suggesting that this is unlikely to explain why no evidence of impact was found.

Table 8: Estimated effect of the Chess in Primary Schools intervention upon children's average maths test scores

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model (intervention; control)	Effect size (95% CI)	p-value
ITT	1,965 (0)	70.0 (67.9 to 72.1)	1,900 (0)	69.2 (66.8 to 71.7)	3,865	0.01 (-0.15 to +0.16)	0.93
CA-ITT	N/A	N/A	N/A	N/A	3,865	0.01 (-0.16 to +0.17)	0.93

Notes: Authors' calculations. Effect size estimates based upon an OLS regression model, controlling for Key Stage 1 maths, reading, writing and science test scores. ITT refer to Intention-To-Treat estimates. CA-ITT refer to the instrumental variable (Contamination Adjusted Intention-To-Treat) results.

Differences in treatment effects by sub-group

Table 9 presents results for three sub-groups: boys, girls, and children who were eligible for FSM. The estimated effect of the intervention on the latter was 0.01 (95% confidence interval running from -0.18 to +0.19). For boys, the impact was -0.02 standard deviations (95% confidence interval running from -0.17 to +0.13) compared to +0.03 for girls (95% confidence interval -0.14 to +0.20). However, a formal test of the gender-by-treatment interaction failed to reject the null hypothesis of no difference between boys and girls at conventional thresholds. Overall, there is little evidence that the intervention had any impact upon the pre-specified sub-groups after one year.

Table 9: Estimated effect of the Chess in Primary Schools intervention upon sub-groups

	Raw means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model	Effect size (95% CI)	p-value
Boys	994 (0)	71.0 (68.8 to 73.2)	931 (0)	71.3 (68.6 to 74.0)	1,925 (0)	-0.02 (-0.17 to +0.13)	0.77
Girls	971 (0)	69.0 (66.7 to 71.4)	969 (0)	67.2 (64.6 to 69.9)	1,940 (0)	+0.03 (-0.14 to +0.20)	0.73
FSM	641 (0)	65.6 (62.6 to 68.5)	680 (0)	64.8 (61.8 to 67.7)	1,321 (0)	0.01 (-0.18 to +0.19)	0.95

Notes: Authors' calculations. Effect size estimates based upon an OLS regression model, controlling for Key Stage 1 maths, reading, writing and science test scores.

Secondary outcomes

Mathematics sub-domains

Table 10 provides the estimated impact of the treatment on each of the maths sub-domains (paper A, paper B, and mental arithmetic). The effect size is very close to 0 on each occasion. This further supports the finding that the intervention had no impact upon maths achievement after one year.

Table 10: Estimated effect of the Chess in Primary Schools intervention upon different components of the Key Stage 2 maths test

	Raw means					Effect size	
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model	Effect size (95% CI)	p-value
Paper A	1,965 (0)	27.5 (26.7 to 28.3)	1,900 (0)	27.2 (26.1 to 28.2)	3,865 (0)	0.01 (-0.15 to +0.16)	0.91
Paper B	1,965 (0)	28.9 (28.0 to 29.9)	1,900 (0)	28.7 (27.6 to 29.7)	3,865 (0)	0.00 (-0.16 to +0.17)	0.96
Mental arithmetic	1,965 (0)	13.6 (13.1 to 14.0)	1,900 (0)	13.4 (12.9 to 13.9)	3,865 (0)	0.00 (-0.12 to +0.13)	0.94

Notes: Authors' calculations. Estimates based upon an OLS regression model, and are based upon ITT.

Reading and science

Table 11 turns to examine spillover effects into two other academic subjects: reading and science. The point estimate was -0.06 standard deviations for the impact on reading (95% confidence interval from -0.21 to +0.09) and -0.01 for science (95% confidence interval from -0.12 to +0.09). There is hence no evidence the intervention had any spillover impact upon these other subject areas.

Table 11: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 reading and science test scores

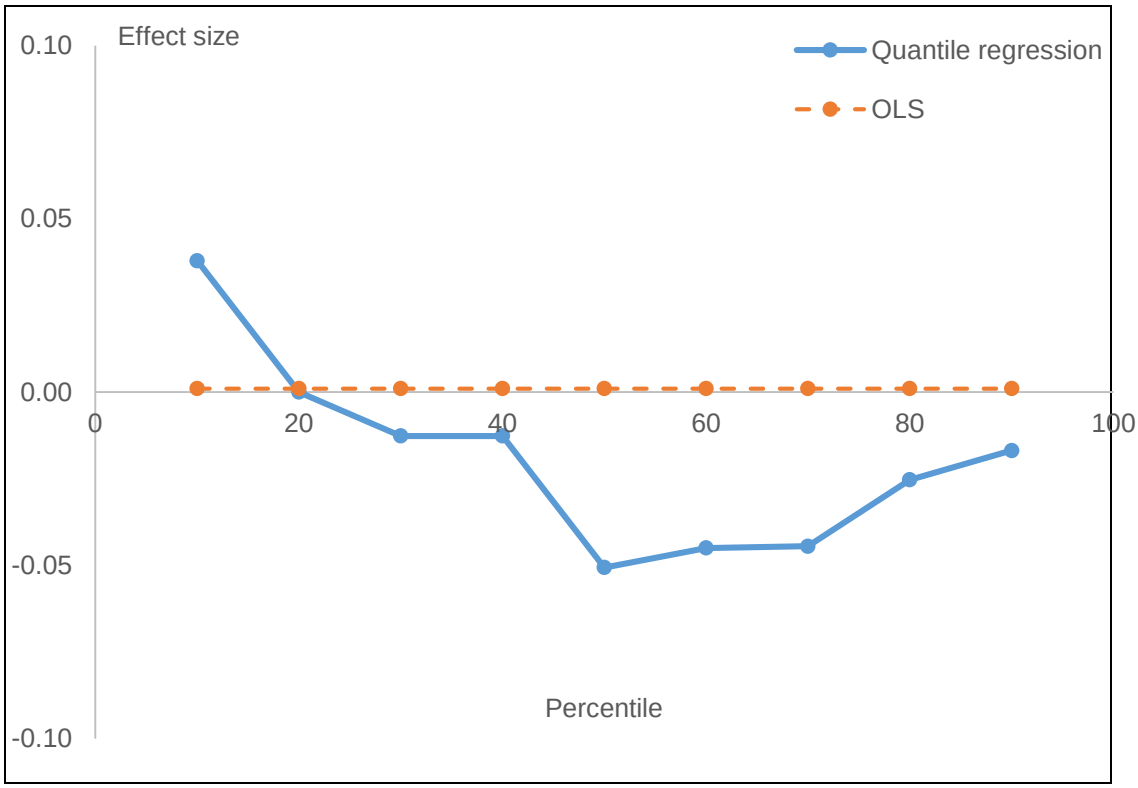
	Raw means					Effect size	
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	n in model	Effect size (95% CI)	p-value
Reading	1,954 (11)	29.8 (29.0 to 30.6)	1,891 (9)	30.1 (29.0 to 31.2)	3,845	-0.06 (-0.21 to +0.09)	0.44
Science	1,965 (0)	4.20 (4.16 to 4.24)	1,900 (0)	4.20 (4.14 to 4.25)	3,865	-0.01 (-0.12 to +0.09)	0.82

Notes: Authors' calculations. Estimates based upon an OLS regression model, and are based upon ITT.

Quantile regression estimates (heterogeneous effects)

It could be that our finding of zero impact upon mean mathematics scores is driven by a large positive impact upon one group (e.g. low maths achievers) and a large negative impact upon another (e.g. high maths achievers). Consequently, Figure 3 presents quantile regression estimates of treatment effect at each decile of the post-test (Key Stage 2) distribution, in order to examine whether the Chess in Primary Schools programme had a different impact upon high and low academic achievers. Running along the x-axis is the percentile of the post-test score distribution where the quantile regression is estimated. The y-axis provides the estimated treatment effect. There is very little evidence that the intervention had any positive effect on either high or low maths achievers. Indeed, many of the point estimates are actually negative, though none are significantly different from zero at even the 10% level. Again, this further strengthens the evidence that the intervention had little medium-term impact upon pupils' maths achievement.

Figure 3: Quantile regression estimates of the impact of the Chess in Primary Schools intervention



Source: Authors' calculations. Solid line refers to the quantile regression estimates. Dashed line provides OLS estimates for comparison. Key Stage 2 total mathematics scores is the dependent variable.

'On-Treatment' analysis

Table 12 presents results from our on-treatment analysis, focusing upon whether the effectiveness of the intervention varied by chess tutor 'quality'. (Recall that 'tutor quality' has been defined using the proportion of children who reported that they liked the chess lessons run by the tutor 'a lot'.) All figures refer to differences in Key Stage 2 maths test scores (presented in terms of an effect size) relative to the control group. There is no clear evidence that children taught chess by tutors of higher quality achieved significantly higher KS2 test scores. Children taught by 'low quality' tutors achieved KS2 test scores slightly below the control group (-0.05 standard deviations) while children with 'medium quality' tutors scored a little higher than the control group (+0.11 standard deviations). However, there is no clear pattern of a 'dose-response' relationship, as the effect of having a high quality tutor was essentially zero. Moreover, none of the estimates presented in Table 12 reach statistical significance at conventional levels. Overall, there is no evidence that the effect of the Chess in Primary Schools intervention varied significantly by whether children liked a particular tutor's chess lessons.

Table 12: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 mathematics scores, by chess tutor quality

		Effect size	
Outcome	**n in model**	**Effect size (95% CI)**	**p-value**
'low quality'	89 schools 3,498 pupils	-0.05 (-0.26 to +0.15)	0.63
'medium quality'	89 schools 3,498 pupils	+0.11 (-0.07 to +0.29)	0.25
'high quality'	89 schools 3,498 pupils	0.00 (-0.27 to +0.26)	0.99

Table 13 presents analogous results for whether the regular class teacher attended the CSC training workshop, as per the study protocol. All figures refer to differences compared to the control group, expressed in terms of an effect size. There is no evidence that the effect of the intervention varied by whether the regular class teacher attended the CSC workshop. For instance, children in treatment schools whose teacher did attend the workshop scored just 0.01 standard deviations higher on their KS2 maths test than children in the control group. This difference is very small and statistically insignificant at conventional levels. Likewise, the test scores of children in treatment schools where the class teacher 'did attend' is little different from the test scores of children where the class teacher 'did not attend'.

Table 13: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 mathematics scores, by whether the class teacher attended the CSC training workshop

		Effect size	
Outcome	**n in model**	**Effect size (95% CI)**	**p-value**
Did not attend	95 schools 3,714 pupils	+0.04 (-0.14 to +0.21)	0.69
Did attend	95 schools 3,714 pupils	+0.01 (-0.16 to +0.18)	0.92

Cost

If a primary school were to participate in the Chess in Primary Schools programme next year, they would incur two main costs. The first is that they would have to pay an annual contribution to the CSC charity. This varies by the size of the school, and whether an after-school or lunchtime chess club is also set up[9]. For instance, a typical two-form primary school with an after-school chess club would be asked to pay £1,800 per year.

The second main cost to schools is that the regular class teacher is expected to complete a one-day training course organised by Chess in Schools and Communities. For instance, for a two-form entry school which needs to pay for supply cover, we estimate this to require a one-off cost of around £400 (assuming a figure of £200 per day for each supply teacher).

[9] Whether a school sets up a lunchtime or an after-school chess club is optional. It is included in the cost estimate presented here, as the after-school club formed part of the intervention evaluated.

Following EEF guidance, we spread this cost over three years, to give an annual figure of £133.

In Table 14, we add these two costs together, and illustrate how the total cost varies by size of school. For instance, we estimate the average annual cost of a primary school to be £1,933 for a two-form entry school. This estimate of the total cost is then divided by the number of pupils (assuming 30 pupils per school form) to provide a cost per pupil. This varies from £52 per pupil in single-form entry schools to £22 per pupil for schools with four forms or more.

It should be noted that schools that participated in this evaluation were not expected to make a contribution to the Chess in Schools and Communities charity during the intervention year; rather, this was covered directly by the EEF grant.

Table 14: Cost to schools to participate in the Chess in Primary Schools programme (costs per year based on delivery over 3 years)

Number of forms	Class lessons	After school club	Teacher training	Total cost	Number of pupils	Cost per pupil
1	£900	£600	£67	£1,567	30	£52
2	£1,200	£600	£133	£1,933	60	£32
3	£1,500	£600	£200	£2,300	90	£26
4	£1,800	£600	£267	£2,667	120	£22
5	£2,400	£600	£333	£3,333	150	£22
6	£3,000	£600	£400	£4,000	180	£22

Process evaluation

Introduction

This section of the report covers the key findings of the process evaluation of the Chess in Primary Schools programme. The process evaluation aimed to explore aspects of the study that provide insight into effectiveness as well as issues such as perceptions of impact and potential improvements and sustainability of the programme.

This section covers:

- Overview of data sources and response rates
- Logic model—developer's' view of necessary conditions
- Implementation
    - Dosage
    - Fidelity to the model
    - Response to the intervention
    - Factors influencing implementation
- Perceptions of programme impact
- Sustainability of the programme
- Lessons for future implementation
- Control group activity.

Overview of data sources and response rates

Table 15 summarises the data sources and response rates. While some individual types of data were more complete than others, across the various sources a good picture of the key themes across the treatment schools has been achieved.

Table 15: Process evaluation data collection—methods and response rates

Method	Sample size	Response – number (rate)
Head teacher baseline survey	100 schools	78 (78%)
Class teacher post-intervention survey – treatment schools	44 treatment schools	28 (64%) schools – 36 teachers submitted data
Class teacher post-intervention survey – control schools	49 control schools	10 schools (20%) – 15 teachers submitted data
CSC tutor survey	24 tutors	23 (96%)
Head teacher treatment schools – post-intervention survey	44 headteachers	18 (41%)
Pupil post-intervention survey – paper	75 classes across 44 treatment schools	776 pupils from 36 classes (48%) across 26 treatment schools (60%)
Treatment arm: observations and interviews	**Numbers conducted**	
Researcher observation of CSC training for study school teachers	2	
Researcher observation of CSC training events for tutors	3	
Researcher observation of intervention delivery	8 classes in 4 schools	
Teacher interviews – telephone or	10 (5 telephone; 5 face-to-face)	

face–to-face following observation	
Headteacher interviews – telephone or face-to-face in observation schools	3 (2 telephone; 1 face-to-face)
CSC head office team – interview	1 face-to-face

Logic model—developers' view of necessary conditions

The logic model (Figure 4) reflects the views of CSC experts on the necessary conditions (inputs and processes in school) for their intervention to be successful. The following sections will reflect on whether these conditions have been met.

Figure 4: Chess in Primary Schools logic model*

Inputs	Processes in school	Change mechanisms (*)	Pupil intermediate impacts	Outcomes
• Sufficient chess boards/pieces • Chess tutor—enthusiastic, reasonable player, good teacher • Team teaching with class teacher/assistant • Good classroom environment, e.g. sufficient space for pairs to play games of chess • CSC primary school curriculum plus CSC work book • One-day basic training course for tutors • One-day basic training course for teachers • Whiteboard or manual display board • Additional training for tutors	Learning chess 1 hour a week for 30 weeks as part of the school curriculum. A CSC tutor using the CSC curriculum for primary schools teaches the chess. The class teacher/assistant plays an active part in the chess lesson.	***Education*** Children learn chess—chess playing requires skills that support maths and English ability. CSC tutor teaches using a graded chess curriculum—lessons are fun and interactive. Tutor (with teacher/assistant support) differentiates teaching to meet different learning needs of individuals within a class. Pupils play chess games together and learn new things about each other. ***Persuasion (or encouragement); modelling*** Pupils who do not excel academically can show an aptitude for chess and gain confidence/recognition. Teacher/assistant learns/plays chess with the children. Tutor promotes and models positive attitudes and behaviour in the context of teaching chess rules and etiquette, e.g. silence, concentration.	• Improved concentration and perseverance • Improved logical thinking and problem solving • Improved confidence and self-esteem • Improved behaviour in school • Improved communication of complex ideas • Excelling of those with particular needs, e.g. gifted and talented, special educational needs, more solitary, etc. • Ability to play a reasonable game of chess • Improved creativity	• Improved KS2 maths SATs results • Improved KS2 English SATs results

**Developed through a Delphi process with CSC experts. This reflects their opinions of what is required to achieve the desired outcomes.*

Implementation

This section describes what was actually delivered in the treatment schools. It considers the dosage received, the fidelity to the model, the response to the intervention and key factors that affected implementation.

Dosage

Six treatment schools chose not to participate in the intervention. The headteachers for three of these six schools returned baseline survey data, prior to randomisation. At that stage two of these three school leaders had been very keen and one fairly keen on taking part in the study. Information provided by CSC suggested that the main reasons for the subsequent withdrawals were practical changes that impacted on the feasibility for the school of being involved, such as turnover of key staff.

The chess lessons were delivered in all 75 Year 5 classes in the 44 treatment schools which delivered the intervention. The one-hour chess lessons ran, in all participating schools, from midway through the autumn term (2013) to near the end of the summer term (2014). According to the tutor and teacher data, only one third of schools received the full intended dose of 30 hours—see Table 16.

Table 16: Amount of chess teaching delivered in a school

	30 hours	25–29 hours	Less than 25 hours
Number of schools	14 (33%)	26 (59%)	4 (9%)

The main reason for this shortfall in intended hours of delivery was that the classes commenced midway through the first term and not at the start. The cause of this delay was the completion of trial processes by CSC with schools. If there were few or no rescheduling requests by the school, then achieving the 30 hours was possible. But where chess lessons were occasionally cancelled by schools, which was not uncommon, completing the full 30 hours became difficult to achieve.

Reach

Teacher and tutor survey data shows that the programme reached its intended recipients in the 44 participating schools. While a few teachers reported via the survey that a few children did not routinely participate in the chess lessons, this was unusual. These children were removed from the lessons either due to other demands on their time (such as instrument lessons) or being perceived by the teacher not to be managing the chess.

Fidelity to the model

In general, intervention delivery adhered fairly closely to the intended programme, as spelled out in the logic model, though there were some key areas where expected inputs and processes of delivery were not completely as intended and/or where there was variability across the sites. Table 17 summarises the adherence and deviation.

Equipment and curriculum

There was clear fidelity to the model in terms of equipment, facilities and the curriculum. Classrooms and equipment were deemed fit for purpose by tutors and the chess sets provided by CSC at the start of the intervention remained available throughout the year. All tutors used the CSC curriculum. Children were given worksheets to complete, as part of the curriculum and work books were introduced as an additional resource by most tutors. As intended by the developers, the tutors did make adaptations to the curriculum. Examples of adaptations included running a class competition over several lessons and using video clips of grand masters playing in a competition. While some

tutors were using the curriculum in digitalized form via whiteboards or TV screens, many were not. Those that weren't were using the manual display board provided by CSC.

Tutor background and skills

The Chess in Primary Schools programme was to be delivered by a chess tutor who was to be 'enthusiastic, a reasonable player and a good teacher'. Twenty-three CSC tutors taught the intervention across the participating schools; there was considerable continuity with very little turnover of tutors in schools. The tutors' survey provided the following information about their background and skills:

- All the tutors were proficient chess players, mainly describing themselves as average or strong club players.
- Just under half the tutors had some kind of teaching qualification, including one who was a qualified primary school teacher, five who had secondary/adult education teaching qualifications and six a specialist chess teaching qualification.
- Five tutors started working for CSC in the year of the study; ten had worked for CSC for 1–2 years and eight for 3 or more years.
- Before the study, 18 tutors had had a moderate amount or a lot of experience of teaching chess to primary school children (in clubs, schools, etc.) while 4 had no or a little experience (1 non response). Of the 23 tutors, 18 had been on the one-day CSC training for tutors, teachers and prospective tutors, which covered the CSC curriculum. During the intervention year, 17 of the tutors attended additional weekend training seminars, a new initiative organised by CSC to support and develop their teaching and classroom management skills.

However, while teachers and headteachers were generally very positive in interviews and surveys about many aspects of the tutors' input (including their enthusiasm), in many schools concern was raised, to varying degrees, about the tutors' teaching and classroom management skills. This theme also emerges from the pupil survey. This will be discussed further in the section 'Factors affecting implementation'.

Tutor delivery in the schools

Individual tutors usually worked in one or two treatment schools (range 1–6 schools). Where a school had two or more Year 5 classes the same tutor usually worked with all the classes, generally in consecutive lessons. Variation to the original planned intervention came with the introduction of paired tutors delivering lessons in some schools. Tutor survey data indicated that pairs of tutors worked in 6 schools for the whole intervention period and in 15 other schools for part of the intervention period. CSC staff confirmed that these pairings generally occurred with relatively new CSC tutors or when the intervention was perceived to need a boost. Pairing of tutors is unusual for CSC but was possible for the trial due to the additional resource available.

Teacher involvement

The class teachers were to attend the CSC one-day training course and then be actively involved in the chess lessons, ideally team teaching with the tutor. In fact, only 31% of teachers attended the training. Furthermore, while most teachers did engage with the lessons their engagement was less extensive than hoped for by many tutors. These deviations are discussed in more detail in the section 'Factors affecting implementation'.

Lesson replacement

Class teacher survey data received from teachers in 30 schools (68% of the treatment group) showed that the chess lesson most commonly replaced a 'topic'/humanities lesson; others replaced included music or PE. However, seven from this group replaced a maths lesson—six wholly, and one

partially—and one school said they replaced an English lesson for the whole of the intervention year. The replacement of a maths or English lesson with the chess lesson was a clear departure from the intended programme. The reasons for this replacement are not completely clear, though many teachers and headteachers reported difficulties fitting all aspects of the curriculum into the Year 5 timetable. Additionally, there appeared to have been some misunderstanding by, and within, schools of what was expected of them for the Chess in Primary Schools programme. This occurred despite contact from CSC with all treatment schools, when it emerged during the first few months of the programme that some schools were replacing maths lessons.

Table 17: Intervention delivery—achievements and variations from plan

Achieved as planned	Variation from plan
Inputs • Sufficient chess boards and pieces • Chess tutor—enthusiastic, reasonable player • Good classroom environment • CSC primary school curriculum plus CSC work book • One-day basic training course for tutors • Whiteboard or manual display board • Additional training for tutors *Processes in school* • A CSC tutor using the CSC curriculum for primary schools teaches the chess	*Inputs* • Team teaching—*tutor/teacher planned collaboration often limited* • Chess tutor good teacher—*some lacked class differentiation skills* • One-day basic training course for teachers—*very low uptake* *Processes in school* • Learning chess 1 hour a week for 30 weeks as part of the school curriculum—*most schools achieved 25–29 hours* • Replacing a lesson other than maths—*at least 7 schools replaced maths lessons* • The class teacher/assistant plays an active part in the chess lesson—*happened in some but not all classes*

Response to the intervention

This section considers the views of the key stakeholder groups—pupils, teachers, headteachers— on the acceptability of the Chess in Primary Schools programme. A limitation of this part of the report is that it is based upon only 774 responses from the initial intervention sample of 1,900 (41%).

Pupils—response to the intervention

Pupils were asked in the survey: *How much did you like the chess lessons you had in your Year 5 class that were taught by a chess tutor?* Table 18 shows that very few disliked the lessons, with the majority liking them 'a lot'.

Table 18: How much did children like the Chess in Primary Schools lessons?

Liked the chess lessons	n=774
A lot	53%
A little	39%
Didn't like them	8%

Prior to the chess lessons, over half (57%) of the pupils said they had never played chess before; whereas less than a fifth (17%) said they had played at least weekly before the chess lessons. The proportion of children who liked the chess lessons 'a lot' was significantly higher for those who had regularly played chess (72%) than those who had not played before the Chess in Primary Schools intervention began (44%, $p<.01$).

Children from schools in the study with the historically lowest attainment scores were significantly more likely to dislike the chess lessons ($p<.01$). More unexpectedly this also appeared to be the case in schools with the relatively lower FSM proportions ($p<.01$). One reason for this appears to be that more pupils in relatively less deprived communities had had more previous access to chess and as a result some of these children reported finding the lessons too slow and insufficiently engaging.

Tutor effects on how much pupils liked the chess lessons

Pupil data was received from 19 of the 23 tutors. A tutor effect emerged when considering the degree of liking lessons. Proportions of pupils who liked the lessons 'a lot' ranged across tutors from 15% to 76%. Five tutors had more than two-thirds of the pupils they taught like the chess lessons 'a lot', whereas five tutors had less than a third. No key background features (e.g. teaching qualifications, years working for CSC) varied significantly between those tutors where a greater proportion of children liked the sessions a lot and those where they did not.

When extent to which pupils liked the chess lessons was analysed by school, it became apparent that there was one school where there was a significant number of children who disliked the chess lessons (n=14 or 50% of the pupils who responded from this school). Process evaluation interviews indicated that the teacher of this class had been dissatisfied with the ability of the tutor to keep children of different abilities engaged with the lessons.

What pupils liked best about the chess lessons

Pupils were asked to write free text answers to the question: *What did you like best about the chess lessons?* Common responses were:

- **Playing chess games and/or mini games**. This was the most frequent answer to this question. Many who said this referred in particular to the fun of playing their friends/classmates. This theme was strongly reinforced by teachers who gave many examples, in interviews and surveys, of pupils choosing to play chess with their friends, rather than other activities available to them during free time in school.
- **Being taught the theory of the game**—such as the rules for the different chess pieces or strategy. This category of response implied a satisfaction with how these theoretical concepts were explained by the tutor.
- Particular **tutor attributes** such as being funny or having specific skills.
- **Learning a new skill** that was both fun and challenging.
- **Choosing friends to play against.** Many said that they disliked having opponents (who were not their friends) chosen for them by the tutor. However, there were children who prioritised having an opponent with whom they would have a good and challenging game over playing particular friends. These pupils welcomed tutors helping organise this.

'I liked that every lesson we had was clear and understandable.' (Year 5 pupil)

'I liked chess in Year 5 because it helps your brain to think and it is good for knowledge. I also like it because you play with your friends.' (Year 5 pupil)

Interestingly pupils rarely mentioned that they liked winning. This suggests that the tutors had succeeded in placing the emphasis on gaining satisfaction from the process of learning and playing rather than securing a victory at the end of a game.

What pupils liked least about the chess lessons

Pupils were also asked what they liked least about the chess lessons. Common themes were*:*

- **Too much talking by the tutor**. This was the most frequent response to this question. Common sub-themes to this response were too much: repetition of what they had learnt in a previous chess lesson; time on the carpet at the beginning of the lesson; talking by the tutor at the expense of time for playing games; interrupting of their games by tutors.
- **Losing**. For some it was evident that they lost frequently and unsurprisingly they disliked this.
- Finding the lessons either **too easy or too hard**.
- **Disliking the rules** set by the tutor, including the need to play quietly and not being allowed to choose who they played with, or conversely wanting stricter discipline to control the disruptive behaviour of their peers, during the lesson.
- **Not been chosen** as anyone's partner or generally finding the partnering process stressful.

'I felt that the teacher spends most of the time explaining things that he had already told us in the previous lesson which means we hardly get to play/finish our game.' (Year 5 pupil)

The pupil acceptability data, both quantitative and qualitative, suggested that most children liked the chess lessons, but it also illustrated the extent of the challenge for tutors because of the diversity of pupil need within classes and the need to balance fun with applying clear rules.

Headteachers and teachers—response to the intervention

Teachers and/or headteachers from 39 treatment schools provided data for one or more of the different data collection activities aimed at them.

Teacher views before the intervention

Prior to randomisation, the baseline survey of headteachers showed that these school leaders were generally very enthusiastic about the Chess in Primary Schools intervention, regardless of personal chess-playing experience or level of chess activity in the school in the past five years. Any reservations were generally about the difficulties fitting the lessons into the timetable.

While the majority of class teachers said they were initially keen to try the intervention, there were some who admitted—retrospectively in the post-intervention teachers survey—to some initial ambivalence (29%) or reluctance (14%). Initial concerns were that the chess lessons would be difficult to fit in to a tight Year 5 timetable, and that low-achieving children would lose out most, through loss of core subject time and difficulties engaging with learning chess.

Teacher views during and after the intervention

The class teachers who engaged in interviews and observations and/or submitted post-intervention survey data were, in general, very positive about many aspects of the programme. Most of them rated the performance of the CSC tutors in their school as 'good'. Many teachers commented positively on particular qualities of the tutors—particularly their enthusiasm for chess and the impact this had on enthusing and engaging the children as well as their positive interactions with the pupils.

'[The tutor] has been brilliant for us, the manner in which he delivers the lesson, his engagement with the children and his ability to calmly solve their game problems has been great. The children look forward to seeing him each week.' (Class teacher)

Examples of high quality teaching practice by the tutors and use of innovative resources to complement the CSC curriculum were reported in surveys and interviews and also observed. These included examples such as tutors tracking each child's progress week by week in order to meet their individual learning needs, and use of specialist chess software such as ChessBase on the whiteboard as well as other IT resources.

Headteachers who submitted post-intervention survey data were also generally very positive about the Chess in Primary Schools programme. Some took great interest in the chess lessons, as observed by researchers conducting observations who saw headteachers 'dropping into' chess lessons and giving out prizes for class chess competitions. Others relied on class teachers to update them on progress with the chess lessons.

'We are a chess loving/playing school now! We never even spoke about chess before and I wonder if some of the kids knew what it was! It has been an amazing experience, we love it here!' (Headteacher)

The key criticisms of the programme from teachers centred on the tutors' performance. These are detailed in the 'Factors influencing implementation' section below.

Chess tutors—views on responsiveness

Most tutors also considered that the intervention had gone either very well or well in the classes in which they had taught. In general they were very or fairly satisfied with the support they received from headteachers and teachers. Furthermore, most tutors felt well supported by CSC and were very positive about the CSC curriculum. Only one tutor said the lessons had gone badly in one class though many tutors thought there could have been improvements in the level of engagement of the class teacher. This issue, which was the tutor's key criticism, is discussed in the 'Factors influencing implementation' section below.

Factors influencing implementation

The process evaluation data suggests that there were two key factors that inhibited implementation of the Chess in Primary Schools programme in treatment schools: the tutors' teaching skills and the level of involvement of class teachers.

Implementation inhibitor: tutors' teaching and class management skills

The most common criticisms of the programme by teachers and headteachers in interviews and surveys were in relation to tutor teaching and class behaviour management skills. Over half the tutors had limited teaching training but there was no expectation by school staff that tutors should be trained teachers, in fact their status as experts in chess was considered by teachers to be key to gaining respect from the pupils. However, the ability of tutors to engage all the pupils and control the class was a clear area of concern.

'My class is very difficult in behaviour and you need to be firm and strong to maintain order. At times this was lacking [from the CSC tutor] and I had to take over to control the class.' (Class teacher)

Sub themes from teacher data on this issue included the following:

- Tutors stood at the front of the class and talked too much, particularly at the start of the lesson. Pupil data supported this.
- Where little or no technology was used in the teaching this increased the risk of engagement problems—this was supported by the observation data.

Many of the tutors acknowledged, in the survey, that their biggest challenge was accommodation of the wide diversity of pupil academic ability in most classes and many wanted more support with this from the class teacher. Conversely teachers expressed the view that it was important to the success of the lessons for the pupils to perceive the tutor, as opposed to the class teacher, as being in control of the lesson and therefore teachers were reluctant to intervene too readily. One tutor, who had worked for CSC prior to the study, commented on the more challenging nature of the treatment school

classes with the suggestion that the range of ability in study school classes presented a particular challenge.

A number of teachers said that tutors needed more training in this aspect of their role.

'Professional development for tutors would be good on how to keep as many of the children engaged at any one time as possible and how to make the sessions as pacey as possible…that is the key to a successful lesson.' (Class teacher)

CSC offered a one-day training course and a number of seminars for tutors to attend. Tutors raised concerns about the usefulness of the one-day training with 5 of the 18 tutors who had attended stating that they were not satisfied with the course. An intention of this training, which deliberately mixes tutors and classroom teachers, is to achieve 'cross fertilisation' of experiences and views. Many tutors enjoyed this aspect of the training. Some however felt that an important focus for them should be on learning key teaching skills and that this could not be achieved while the group was mixed in this way.

For many teachers and headteachers the tutor teaching and class management issues discussed above were expressed as relatively minor concerns and were offered as constructive suggestions for improvement in the future to a programme that they felt had great potential. In a small number of schools, however, classroom management and delivery issues were perceived to be a more serious problem and made any future engagement with the programme by the school questionable.

It is evident that some teachers who had concerns made efforts to find ways to give tutors feedback and suggestions for change. This was clearly not easy when an expectation of such feedback did not appear to have been discussed. In general, where feedback happened, it appeared to be well received by the tutor and led to positive change in tutor performance.

'I emailed [the tutor] and said that there were too many long periods of listening to him talking and suggested that he should break it down with them having more chances to have a go at things. And he responded to this very well and children's engagement improved a lot more.' (Class teacher)

While most tutors liked the CSC curriculum that they used in the classroom, one tutor was overtly critical of it, saying it was old-fashioned and needed to have more 'hooks' that would capture the pupils' attention. The curriculum workbooks were generally liked by the tutors—*'they facilitate pattern recognition'*— but pupil data suggested mixed views on these workbooks.

Implementation inhibitor: class teachers' role in the lessons

From the tutor and CSC head office staff perspective the solution to the issues of delivery and classroom management was a high level of involvement of the class teacher, including attendance at the one-day CSC training course. As stated previously this involvement was less extensive than hoped. For example:

- Only 23 class teachers from 19 treatment schools (31% of teachers) attended the CSC one-day training. The main reasons given by the teachers were that they weren't aware of the training, lack of time or lack of permission from their manager.
- The majority of class teachers provided some help with the chess lessons but few tutors described a formalised team teaching approach with the class teacher in the lessons. Interview, survey and observation data suggests that one reason for this was differing perceptions between teachers and CSC staff about appropriate approaches to leadership of a lesson.
- In the chess lessons, approximately two thirds of the teachers played or learnt to play chess alongside the children. The other third did not take part in playing. CSC tutors suggested this was due to teachers feeling self-conscious about being seen by their pupils in the role of a learner, as opposed to an expert. While in some classes teaching assistants helpfully played

when a teacher did not, this was not considered by tutors to have as powerful an effect on pupils as observing their teacher engaging in learning a new skill.

There were some examples given by tutors of classes where there was little or no joint working or support of any kind from teachers.

'He [the class teacher] would sometimes be there as I arrived then would leave and just come back at the end of the day to dismiss the class.' (CSC tutor)

The evidence from the process evaluation was that many teachers were not fully aware of these joint-working expectations, even by the end of the intervention. It was clear that in many schools, tutors and class teachers had had little time to liaise, prior to or during, the intervention period.

One aspect of the programme which should have helped with clarifying this role was the one-day CSC training course for teachers and tutors. Uptake of this was low, but those teachers that did attend had mixed views. Most reported enjoying the training, but did not think it was critical to the success of the intervention in their school. Teachers reported that the emphasis on the specifics of chess in the training was too strong (and as non-chess players went over their heads). They wanted a greater focus on how to make the chess lessons as successful as possible in the classroom setting.

Perceptions of programme impact

Teachers were dubious about whether the chess lessons would have impact on the primary outcomes of maths and English attainment. The teachers' survey showed that only about a quarter of teachers thought the chess lessons would have quite a lot of impact on pupils' maths attainment; half thought they would have a little impact and a few thought there would be no impact or were uncertain about impact. For English attainment, teacher views were roughly split between predicting the chess lessons would have a little impact and no impact.

Teachers were, however, overwhelmingly positive that the chess lessons would have impact on pupils':

- thinking/cognitive skills;
- confidence/self-esteem;
- ability to cope with winning/losing;
- concentration; and
- ability to play a game of chess.

'They [pupils] have mostly developed in their ability to slow down and really think about problems.' (Class teacher)

'I think this has been an invaluable experience for the children in my class. It has raised morale, achievements and sportsmanship. It has revealed hidden skills and talents, crossed the barriers between games and education and should, in my personal opinion, be made part of the National Curriculum.' (Class teacher)

While many teachers were also positive about impact of the chess lessons on peer relationships and pupil behaviour, some teachers thought there was no impact (or occasionally negative impact) on these.

Perceived impact on lower achievers

In the survey and in interviews, teachers gave examples of pupils at both ends of the academic ability spectrum that had enjoyed and benefitted from the lessons, with progress by individuals who were at the lower end particularly being selected as examples of positive outcomes of the programme.

'Some of the children that wouldn't have expected to excel have—and others want to pair with them—and this has been really nice for those children.' (Class teacher)

However, there was also considerable concern expressed by teachers and headteachers in some schools that the lower achievers got left behind in the lessons and as a consequence disengaged and often became disruptive. There were also concerns expressed by a few teachers that slower learners might be relatively negatively affected by the loss of a maths lesson (in schools where chess replaced maths).

'In mixed ability groups they [children with additional needs] are not able to access the learning as quickly as others so they become frustrated. The other children become impatient because they want to get on with the game.' (Class teacher)

Impact on the school

The impact of the programme on the school as a whole was mentioned by many schools. The majority of teachers reported on the survey that chess playing in the school had spread beyond the Year 5 chess lessons. Examples given were newly established or reinvigorated chess clubs and chess playing during free time in classrooms and/or the playground.

'Children from Year 2 upwards have benefited from the purchase of a giant chess set in the playground—Year 5 have been able to teach chess to others.' (Headteacher)

Sustainability of the programme

Tutors were fairly optimistic about the potential for chess playing to continue among the pupils they had taught. They reported that in the majority of classes over three-quarters of the children could play a reasonable game of chess by the end of the intervention. They also thought that about two-thirds of schools had a member of staff who was confident enough to teach chess themselves and could carry on doing so within their school.

The teacher survey data also suggested that many teachers intended to continue to incorporate chess in the classroom and/or school. Examples given were setting up a chess club and using some of the maths challenges based around chess. However, teachers were clearly concerned about potential barriers such as their own lack of confidence and time pressures and there was no suggestion that chess lessons of the type delivered by the Chess in Primary Schools programme would continue, unless the school purchased the programme.

'I would definitely encourage schools to teach chess and I would like to become more confident to have a go myself. I don't feel ready to teach it yet but possibly in the future I would.' (Class teacher)

Level of continued chess playing

The pupil survey was carried out approximately seven months after the chess lessons finished. Pupils were asked if they were still playing chess. Table 19 shows the responses to this question.

Table 19: Amount of chess playing by pupils 7 months after the intervention finished

Amount of chess played – 7 months post intervention	% All pupils N=772	% Where not played before Chess in Primary Schools	% Where had played frequently before Chess in Primary Schools
At least once a week	28	17	49
Between 1 and 3 games a month	39	35	39
Not playing any chess at all	34	48	12

- Unsurprisingly, there was a significant difference in continued playing between those who had played chess before the lessons, and those who had not ($p<.01$).
- There was no significant difference in continued playing between children from schools with historically lower achievement or higher proportions of pupils with free school meals status.

Most children who were still playing chess, said they were doing so with family and friends (88%, n=449). Some classes were given chess sets to take home by teachers (for example as a Christmas present) to support them in extending their playing into the home. Additionally, 68 children from 14 schools were playing in a school chess club, and a further 56 pupils were playing more informally at school. Twenty-nine children, including seven who had never played chess before the lessons, said they were now playing in chess clubs outside of school.

Continuation with the Chess in Primary Schools programme

Despite the wide scale acceptability of the programme to schools as a free one-year study intervention, when asked in the survey about whether they would pay for the programme during the following year, most teachers were very uncertain. The main barriers mentioned were cost, pressure on curriculum time, concern about the view of Ofsted and potential adverse effects on groups of learners, especially slower learners.

'It's a whole afternoon out essentially and…has really eaten into the curriculum—if they had all been completely engaged and excited about it that would be different but there has been this group that has struggled with enthusiasm.' (Class teacher)

CSC reported that 24 treatment schools paid for their programme to continue, as part of the curriculum, in the school for a second year. Information on which year groups were receiving the programme in these schools was not available. However, no school was allowed to continue the programme with the pupils who had participated while in Year 5, to keep the intervention dose the same for all participating schools.

Lessons for future implementation

Analysis across sources of process evaluation data provided a range of formative findings to inform future implementation:

1. **A set-up meeting between CSC head office staff and the headteacher** should be part of the 'sign-up' process. This could cover background detail for headteachers on the perceived benefits of learning chess (to help with justification for Ofsted), and expectations of CSC and

of the school, and involve clear terms of reference and agreement with staff. Effective communication of the key points from this meeting from headteacher to class teachers is important. In particular headteachers should convey the expectation that class teachers and tutors team-teach and teachers learn to play chess alongside the children.

2. **A set-up meeting/training event on school premises between the chess tutor and relevant teaching staff** should be a part of the programme. This should be paid time for the tutor and would replace the current one-day CSC training day for teachers. The aims of this meeting would be: to establish a good teacher/tutor working relationship; for teachers to learn about the programme; for teachers and tutors to share expectations and requirements of their respective roles in the classroom; for tutors to learn about the classes receiving the lessons and how best to work with them; to teach non-chess-playing teachers the rudiments of chess.
3. The training provided by CSC for tutors, including the one-day introductory course, should **focus on teaching techniques and class management skills**. It should have at least some input from an experienced teacher/teacher trainer who would provide training on teaching techniques, particularly those that assist with differentiation in the classroom. Training for tutors should also focus on minimising risk to pupils who lose chess games frequently or find aspects of the partnering process stressful.
4. **Opportunities for more development of individual tutors' teaching skills** should be created. These could include: opportunities to work with and be mentored by a CSC tutor who excels as a teacher; putting in place systems that require schools to provide termly feedback, including pupil views, on tutor performance; and constructive suggestions for raising this.
5. **Audits to ensure tutor quality** should be conducted by CSC with associated tutor supervision and support where appropriate to improve performance.
6. **The chess lessons should be more interactive** and CSC tutors should be trained and supported in the use of new technologies.
7. **CSC should consider tutors working as a pair** in the classroom as the ideal model, to be achieved wherever possible. Other approaches that support effective differentiation in the classroom should be considered. Positive examples of splitting a single class or two amalgamated classes, by chess-playing ability once sufficient chess classes have taken place for this to be assessed for each pupil, were reported by teachers. This approach was reliant on a teacher or teaching assistant being able to teach one of the two groups.
8. **The chess lessons should start at the beginning of the academic year** in order that the full 30 lessons can be readily fitted in and should not replace a maths or English lesson.
9. Reverting to the more common CSC practice of **delivering the lessons to a younger year group** would potentially reduce curriculum pressure concerns.

Control group activity

Fifty schools were initially randomised to the control group, but one refused their randomisation status, and booked CSC to deliver chess lessons to their Year 3 students.

Thirty-eight headteachers of schools in the control group completed the pre-randomisation baseline survey. Comparative analysis suggests that there was little difference at baseline between the two arms of the trial in terms of chess-playing activity in their school, with 45% of control schools and 48% of treatment schools saying that this occurred. In both trial arms seven of these chess-playing schools reported only occasional informal chess-playing in the classroom while four schools said there was a lot of chess-playing—including in a chess club and informally in the classroom. Five control schools and six treatment schools said they ran a chess club but did not appear to have any other chess-playing in the school.

The response rate to the follow-up control teachers' survey was extremely low (20%, see Table 13). This was despite concerted efforts by the research team, which included online, telephone and paper-based completion routes. This low response could be the result of their disappointment at their trial

arm allocation. Equally though, this could also be an indication of lack of engagement or even knowledge of the programme by class teachers in schools that never received the intervention.

Our survey results, although limited, did not show any sign of compensation rivalry relating to chess initiatives within the control schools. A third of those that responded continued to offer chess clubs and recreational chess; none had external chess activities. We do not have a complete picture of whether extraordinary additional maths or English support was brought into control classes to compensate for the lack of the Chess in Primary Schools lessons. Four of the 15 control teachers who responded to the questionnaire had some additional maths programme for their Year 5 classes. This level, if replicated across the full set of control schools, would not be considered extraordinary. Unfortunately, the poor response means that we cannot be certain.

Conclusion

Key Conclusions
1. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
2. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
3. There is no evidence that the intervention had a positive impact on Key Stage 2 scores for children eligible for free school meals (FSM).
4. Although a current school teacher is allocated to every chess class, it is desirable for the tutors themselves to have strong class management and teaching skills. Without these, it was difficult to ensure that all children were suitably engaged in the chess lessons.
5. For successful implementation, class teachers need to work closely with the tutor and actively contribute to the intervention. It was felt that classes were less effective if the teacher did not actively take part, or was present only at the beginning and end of the class.
6. Half of the pupils who participated in the trial said that they liked the chess lessons a lot, and only 8% reported that they didn't like them. School teachers were very positive about the intervention and its impact on pupils' skills and behaviour.

Limitations

The findings outlined above should be considered within the context of the limitations of this study. The following factors particularly stand out:

1. Focus on academic achievement. The purpose of this trial was to examine the impact of Chess in Primary Schools upon children's academic achievement. Although we find little evidence of any impact, we cannot rule out the possibility that the programme has wider benefits for children. This includes potential impacts upon their well-being, self-confidence and non-cognitive skills.
2. Small 'dose' of the intervention. Children have been exposed to the Chess in Primary Schools intervention for just one academic year. This may be a relatively small 'dose' of the programme. A longer exposure may be needed to have a sustained impact upon educational achievement. Little is currently known about the cumulative impact of playing chess over a sustained period of time.
3. External validity. A strength of this RCT is that we have examined external validity, and considered how well the participants compared to the population eligible to receive the intervention. However, the population of interest was quite specific, and had different characteristics from children in England as a whole. It therefore remains unknown how far our results generalise to the rest of the country.

Interpretation

The central hypothesis of this study was that teaching primary school children how to play chess would have a positive impact upon their educational achievement (measured one year after the intervention had finished). This RCT provided very little evidence in support of this hypothesis—the estimated effect on reading, science and multiple elements of mathematics after one year was

essentially zero. This is in contrast to the only other large-scale RCT of the impact of chess on educational attainment that we are aware of, by Boruch and Romano (2011), who detected a substantial effect of more than 0.3 standard deviations for primary school children in Italy. Our results are also in contrast to another recent quasi-experimental study by Gumede and Rosholm (2015), who found a positive effect of chess on primary school children's achievement in Denmark (effect size 0.15).

There are several possible explanations for this difference in results. First, our study was concerned with whether teaching children how to play chess had a *medium-term* impact upon their educational achievement (measured one year after the intervention had finished). In contrast, Boruch and Romano (2011) investigated the immediate impact, straight after the trial had finished. Consequently, their results are more likely to be subject to Hawthorne effects than ours. It is also possible that interventions of this nature have a short-term but not a medium-term impact on academic outcomes, which would explain why an impact was found in the 2011 study but not in this one. Second, our study has used high stakes, external tests as the outcome measure. This is in contrast to Boruch and Romano (2011), and indeed many other RCTs, where the use of low-stakes tests is common. It is possible that the treatment group will be more motivated than the control group when completing such low-stakes tests. Consequently, the study by Boruch and Romano (2011) may have actually been driven by a 'test motivation' effect. Third, the studies were conducted in very different settings. Although Boruch and Romano (2011) did not comment upon the external validity of their study, different findings in the UK should not be unexpected. Finally, we note that 7 of the 44 schools that delivered the intervention chose to deliver chess in place of a maths lesson. However, this is a small proportion of all participating schools, and our robustness checks indicate that this is unlikely to have an impact upon our substantive conclusion.

Security of findings

The Education Endowment Foundation has designed a range of criteria to assess the security of research findings (available from http://educationendowmentfoundation.org.uk/uploads/pdf/Classifying_the_security_of_EEF_findings_FINAL.pdf). Independent peer reviewers are asked to rate each evaluation against five criteria (planned design, power, attrition, balance, and threats to validity). In Table 20 the evaluators present a summary of key pieces of evidence related to these criteria.

Table 20: Evidence regarding the security of research findings

Criteria	Evidence
Planned design	Randomised controlled trial
Power	100 clusters
	Minimum detectable effect from (apriori) power calculation ≈ 0.18
	Actual minimum detectable effect = 0.16
Attrition	0 (0%) of 100 clusters lost due to attrition
	144 (4%) of 4,009 pupils lost due to attrition
Balance	0.05 standard deviation difference in KS1 APS between treatment and control at baseline. Minimal difference in KS1 maths test score distribution.
	% FSM. Control = 36%. Treatment = 33%
Threats to validity	7% of clusters suffer from potential contamination. Robustness of findings tested by conducting a CA-ITT analysis.
	Randomisation, analysis and testing all conducted blind to treatment
	Key Stage 2 tests are high stakes, externally marked and non-specific to the intervention

Other markers	Long-term follow-up built into trial design via NPD
	Protocol published online
	Trial registered with independent organisation
	External validity / representativeness considered
	Randomisation conducted by independent evaluator

Future research and publications

We believe that this study has provided strong evidence that teaching primary school children how to play chess has little lasting impact upon their educational achievement. Future work should therefore concentrate on the potential wider benefits of chess, such as children's well-being and non-cognitive skills. The project team are planning to publish this study as an academic working paper and journal publication in 2016.

References

Aciego, R., García, L. and Betancort, M. (2012) 'The benefits of chess for the intellectual and social-emotional enrichment in schoolchildren'. *The Spanish Journal of Psychology*, 15: 551–9.

Boruch, R. and Romano, B. (2011) 'Does playing chess improve math learning? Promising (and inexpensive) results from Italy.' Available at: http://www.europechesspromotion.org/upload/pagine/doc/SAM_research_synthesis.pdf (accessed 11 December 2015)

Gumede, K. and Rosholm, M. (2015) 'Your move. The effect of chess on mathematics test scores.' IZA DP No. 9370. Available at http://ftp.iza.org/dp9370.pdf (accessed 8 October 2015)

Sussman, J. and Hayward, R. (2010) 'Using instrumental variables to adjust for treatment contamination in randomised controlled trials'. *British Medical Journal*, 340: c2073 doi: 10.1136/bmj.c207

Appendix A. School consent form to access the National Pupil Database (NPD)

"Chess in Schools and Communities" (CSC) programme

National Pupil Database (NPD) agreement form

This form is to be returned to Malcolm Pein, Programme Coordinator, by <INSERT DATE>.

As a school taking part in the "CSC" programme you agree to (i) provide some key information on pupils within your school, (ii) provide consent for the evaluation team at the Institute of Education to access pupils school records held on the National Pupil Database (NPD) and (iii) for the Institute of Education to link the test score data to any additional information collected through questionnaires as part of the CSC programme

The independent evaluation carried out by the **Institute of Education** requires this information in order to conduct a statistically robust evaluation of the CSC programme. Pupils' test scores and any other pupil data will be treated with the strictest confidence. Named data will be matched with the National Pupil Database and shared with the Institute of Education and EEF for research purposes. No individual school or pupil will be identified in any report arising from the research.

I understand and agree that:

- *The school consents to the use of National Pupil Database pupil data for purposes of this evaluation.*
- *That any data collected as part of the evaluation can be matched to individual NPD records, and that this data can be shared with the Institute of Education and Education Endowment Fund for research purposes (at a level of Tier 1 access).*
- *That the school will complete the attached spreadsheet capturing key information on year 5 pupils and send it (electronically) to Malcolm Pein by <INSERT DATE>,*

Headteacher name:			
Headteacher signature:		Date:	

If you have any queries about the evaluation, please contact John Jerrim at the IoE at **J.Jerrim@ioe.ac.uk** or 07590761755.

Any queries relating to the CSC programme can be directed to Malcolm Pein, Programme Coordinator, at <INSERT *EMAIL ADDRESS> or* <INSERT *TELEPHONE>*

Appendix B. Questions CSC were asked to estimate costs

Question 1. Please could you provide an estimate of the average cost of the equipment needed to run the programme per school (e.g. Chess sets etc)

Question 2. How many hours, in total, did the regular class teacher have to attend the training in the CSC programme?

Question 3. How many hours training did the CSC tutors complete?

Question 4. What is the average hourly pay of the CSC tutors?

Question 5. What expenses do you pay the CSC tutors?

Question 6. Do you pay your tutors anything for 'preparation time'? If so, how much? And how many hours (on average) do they spend preparing per class?

Question 7. If a school wanted to take part in your programme next academic year, how much would you charge them?

Appendix C: Padlock rating

Figure 1: Summary grid of criteria for rating the security of evaluation findings

Criteria for interim rating			Adjust	Rating
Design	**Power**	**Attrition**		
Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%	Adjustment for Balance [-] Adjustment for threats to internal validity [-]	5
Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%		4
Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%		3
Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%		2
Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	51-50%		1
No comparator	MDES > 0.6	<50%		0

Appendix D: Cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. Cost ratings are awarded using the following criteria.

Cost	Description
£	*Very low:* less than £80 per pupil per year.
£ £	*Low:* up to about £200 per pupil per year.
£ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Education
Endowment
Foundation

The Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP
www.educationendowmentfoundation.org.uk