



Department
for Education

The Early Education Pilot for Two Year Old Children: Age Five Follow-up

Research report

March 2013

**Ruth Maisey, Svetlana Speight & Vicky Marsh
With Dan Philo**

NatCen Social Research

Contents

Table of tables	4
Acknowledgements	6
Summary	7
Key findings	7
Background	7
Findings	8
1 Research questions, data and methods	9
1.1 Policy and research background	9
1.2 Data sources	11
1.2.1 Survey data	11
1.2.2 National Pupil Database	12
1.3 Linking the original survey data to the NPD	14
1.4 Approach to analysis	15
1.4.1 Matching the pilot and comparison children	15
1.4.2 Age at assessment	17
2 Impact at age five	19
2.1 Overall impact	19
2.2 Impact on subgroups	19
2.2.1 Children who attended better quality settings as part of the pilot	20
2.2.2 Children from black or minority ethnic backgrounds	22
2.2.3 Children in other subgroups	23
2.3 Explaining the results	24
2.3.1 Methodological issues	24
2.3.2 Early years provision received at age three to four	25

2.3.3 School experience during the Reception Year	27
3. Conclusions	30
References	32
Appendix A Propensity score model for overall sample	34
Appendix B Additional tables for subgroup analysis	37
Appendix C Matched and unmatched EYFS profile scores and take-up of early education at age three or four	39

Table of tables

Table 1.1 Profile of children's month of birth	18
Table 2.1 Impact of the pilot on EYFS profile scores and take-up of early education	19
Table 2.2 Impact of the pilot on EYFS profile scores and take-up of early education for children who went to better quality settings	21
Table 2.3 Variation across EYFS profile scores and take-up of early education for pilot children attending settings of different quality	22
Table 2.4 Impact of the pilot on EYFS profile scores and take-up of early education for children with a BME background	23
Table 2.5 Type of early education setting attended at age three/four	25
Table 2.6 Ofsted inspection grades for settings attended by children at age three/four	27
Table 2.7 Characteristics of school attended at age five	28
Appendix table A.1 Propensity score model to model the likelihood of being in the pilot group	34
Appendix table B.1 Impact of the pilot on EYFS profile scores and take-up of early education for children with lower cognitive ability at age two	37
Appendix table B.2 Impact of the pilot on EYFS profile scores and take-up of early education for children who lived in the 20% most disadvantaged areas of the country at age two	37
Appendix table B.3 Impact of the pilot on EYFS profile scores and take-up of early education for children who received more than 7.5 hours of early education as part of the pilot	38
Appendix table C.1 Matched and unmatched EYFS profile scores and take-up of early education at age three or four	39
Appendix table C.2 Matched and unmatched EYFS profile scores and take-up of early education for children who went to better quality settings	40
Appendix table C.3 Matched and unmatched EYFS profile scores and take-up of early education for children with lower cognitive ability at age two	40
Appendix table C.4 Matched and unmatched EYFS profile scores and take-up of early education for children with a BME background	41

Appendix table C.5 Matched and unmatched EYFS profile scores and take-up of early education for children who lived in the 20% most disadvantaged areas of the country at age two 41

Appendix table C.6 Matched and unmatched EYFS profile scores and take-up of early education for children who received more than 7.5 hours of early education as part of the pilot 42

Acknowledgements

We would like to thank all the parents who took part in the original evaluation of the Early Education Pilot for Two Year Old Children for their time and permission to undertake this secondary analysis, as well as all the individuals who contributed to the implementation of that research.

We are also grateful to Adele Rowe, Steve Hamilton, Patrick Towgood and Sophie Faber at the Department for Education (DfE) for their support throughout this project and their valuable feedback on the report. We are grateful to Karen Stevens and Eleanor Cotzias at DfE for providing us with datasets for the analysis and their advice on data interpretation.

Views expressed in this report, together with any errors that remain, are those of the authors.

Summary

Key findings

For children who attended the early years education pilot when they were aged two, there is no evidence that overall they had better outcomes at age five, as measured by the Early Years Foundation Stage profile, than children who did not attend the pilot.

The exception is for children in the pilot study who received early years education in high quality settings. They performed somewhat better at age five than those children who attended low or adequate quality settings as part of the pilot.

There is also no evidence that attending early years education as part of the pilot at age two increased the likelihood of those children attending early years education at age three or four. The exception is for children from black and minority ethnic backgrounds who were more likely to attend early years education when they were aged three or four, if they attended as part of the pilot at age two.

Background

The original evaluation of the Early Education Pilot for Two Year Old Children compared the outcomes for children in the pilot when they were aged three with a comparison group of children who were matched on a range of characteristics. The comparison group was designed to be as similar to the pilot group as possible (Smith et al. 2009). Typically the pilot provided 7.5 hours of early education a week. This original evaluation did not find that the pilot was effective in improving children's outcomes overall when compared with the matched comparison group. However, for those children who received early years education in better quality settings, it was effective in terms of children having a larger vocabulary and more positive relationships with their parents (Smith et al. 2009).

This report examines the same children's outcomes when they are aged five to explore if there are any longer-term benefits for the children who took part in the pilot. The outcome data is taken from the Early Years Foundation Stage (EYFS) assessments. These are carried out in schools by teachers in the year that children turn five (Reception Year).

This study also examines if children who took part in the pilot were more likely to take up early education when they were three or four years old. This is an indirect measure of the pilot's benefit as attending high quality early education has a positive effect on children's outcomes (Sylva et al. 2004). Both the EYFS data and the data on take-up of early years education at age three or four were extracted from the National Pupil Database (NPD).

Findings

There was no evidence of an improvement in the EYFS profile scores of the pilot group at age five when compared with children in the matched comparison group. We also compared the scores for five subgroups to see if there was any effect on particular groups of children. These were:

- Those who attended better quality settings as part of the pilot;
- Those from black and minority ethnic (BME) backgrounds;
- Those with lower cognitive ability at age two;
- Those living in the most disadvantaged areas at age two; and
- Those who received more than 7.5 hours of early education as part of the pilot.

There was no evidence that children in any of these subgroups who took part in the pilot scheme achieved higher EYFS profile scores when compared with children in the matched comparison group.

There were, however, some methodological issues, which affected our ability to detect the effect of the pilot on the subgroup of children who attended better quality settings. The original evaluation found that the pilot only had a positive effect on children's outcomes among children who received a place in better quality settings. In this study, very few children attended settings which were assessed as high quality (Infant and Toddler Environment Rating Scale (ITERS) score of 5 or more). This limited the study's ability to pick up the impact of the pilot on this subgroup of children when compared with children in the matched comparison group.

Further analysis showed that children in the pilot who attended high quality settings (a score of 5 or more on ITERS) performed better in their EYFS assessments when compared with children in the pilot who attended low or adequate quality settings (a score of 4 or less).

The study found no evidence that the take-up of early years education at age three or four was different for children in the pilot and comparison group. The exception was for children from black and minority ethnic (BME) backgrounds, where BME children in the pilot group were more likely to receive early years education at this age compared with BME children in the comparison group. Other research has shown that although take-up of early education at ages three and four is very high, take-up is lower among children from a BME background (Speight et al. 2010). The finding here indicates that the pilot helped improve take-up among children from a BME background.

1 Research questions, data and methods

This study builds on an evaluation of the Early Education Pilot for Two Year Old Children carried out by NatCen Social Research and the University of Oxford on behalf of the Department for Children, Schools and Families (now the Department for Education, see Smith et al. 2009). It uses data from the original surveys and the National Pupil Database (NPD) to explore any longer term effects of the pilot. The focus of the analysis is on outcomes from the Early Years Foundation Stage (EYFS) profile (as an indication of school readiness) and on take-up of early education when children were three or four years old. It explores the overall impact of the pilot and the impact for particular subgroups within the population.

1.1 Policy and research background

The original pilot provided early years education to over 13,500 two year olds and ran between 2006 and 2008. The main purpose of the pilot was to improve children's social skills (e.g. their social confidence and independence) as well as their verbal and reasoning ability. Another aim was to have a positive impact on children's parents and wider family (e.g. on the relationship between parents and their children, and on parents' emotional wellbeing).

The detailed policy rationale for the introduction of the pilot was outlined in the Ten Year Strategy for Childcare (HM Treasury 2004). This was strongly influenced by research evidence (e.g. Melhuish 2004, Sylva et al. 2004) which showed that:

- Children's all-round development is boosted by pre-school experience;
- Such experience is particularly beneficial for disadvantaged children;
- Early entry between two and three years of age is related to greater cognitive gains and improved social behaviour;
- Children's cognitive gains are similar whether they attend full-time or part-time;
- The quality of pre-school settings is related to children's outcomes.

In total, 32 local authorities took part in the pilot, and each authority had a great deal of freedom about how they implemented the scheme. As a result, a wide range of different types of families were targeted. For example, some authorities targeted families using broad criteria, such as living in a disadvantaged ward, while others targeted families on much narrower criteria, such as having a child on the Child Protection Register or with special educational needs.

Typically children in the pilot received 7.5 hours of early years education a week for 38 weeks of the year (although a small number of local authorities offered children 9 or 12.5

hours a week). The pilot places were available in a variety of early years settings but all were required to operate the Birth to Three Matters curriculum.¹

The original evaluation for the pilot found that overall it did not improve the cognitive and social development of the children receiving early years education. However, this overall lack of impact disguised the fact that for those children who were in better quality settings there was an impact on children's vocabulary and their relationship with their parents. More generally, the pilot programme appeared to generate more positive views about formal childcare among families. There was also some evidence that this led to higher take-up of early education at age three (Smith et al. 2009).

In 2010, as part of the Fairness Premium, the Coalition Government announced an expansion of early education. From September 2013, 130,000 two year olds from lower income families will be able to access 15 hours of early education a week. In his 2011 Autumn Statement, the Chancellor committed additional funding to extend this to 260,000 two year olds in England from September 2014.

To support this commitment the Department for Education has introduced legislation to place a statutory duty on local authorities to secure early years education places for eligible two year olds (the Local Authority (Duty to Secure Early Years Provision Free of Charge) Regulations 2012).² These regulations state that two year old children from families meeting the eligibility criteria used for free school meals, and looked after children, will be eligible.

The Government's consultation on the proposed eligibility criteria for the second phase of roll-out from September 2014 ended on 15th October 2012 and ministers are currently considering the responses. The Government's proposals were that children will be eligible if:

- their family (in the current system) meets the income/benefits criteria used for free school meals, or receives working tax credits with annual gross earnings of no more than £16,190
- their family receives Universal Credit and has annual gross earnings of no more than £16,190, or
- they are looked after, have high-level special education needs or disabilities, or have left care but are not able to return home.

This research study aims to feed into the development of early education for two year olds from lower income families to be introduced in September 2013.

¹ Birth to Three Matters was a framework for childcare professionals that aimed to support children in their

² Available from <http://www.legislation.gov.uk/ukxi/2012/2488/introduction/made> [accessed on 5 December 2012]

However, it should be noted that the early education places for two year olds from lower income families which are to be introduced in September 2013 differ significantly from the original pilot. Building on the findings from the original evaluation, the Government:

- has created legally defined eligibility criteria targeted at children from lower income families for 15 hours of early education per week rather than 7.5 hours;
- has communicated the ambition that places are offered in provision that is judged by Ofsted to be good or outstanding; and
- has introduced and revised the EYFS for delivery by providers to ensure consistently high quality early years provision across settings.

As such, the findings from this research study cannot be seen to represent the impact of the early education for two year olds to be introduced from September 2013.

1.2 Data sources

1.2.1 Survey data

The original evaluation measured the impact of the pilot through a specially designed longitudinal study of families taking up a pilot place together with a similar longitudinal study in a comparison group of families. Interviews with families took place at two points in time: first when the child was aged two and before (or just after) the child had taken up their place (this provided baseline data for the evaluation), and at age three after the pilot place had come to an end (this follow-up stage provided data on outcomes).

The sample of families for the pilot was collected from local authorities who ran an opt-out process. The contact details of all the parents who had not opted out of the evaluation were then passed to NatGen. The comparison sample was selected from Child Benefit records and focused on disadvantaged areas of England where the pilots were not operating.³ As with the pilot sample, all comparison families were given an opportunity to opt out of the evaluation before being invited to take part in the research.

The baseline interviews among families in the pilot were conducted in two waves – the first from January to March 2007 and the second from April to June 2007. The baseline interviews among families in the comparison group were conducted in three waves – the first from March to April 2007, the second from June to July 2007, and the third from November to December 2007. All the baseline interviews were conducted face-to-face and were supplemented by a short paper self-completion element for parents.

Follow-up interviews were conducted with parents who completed all elements of the baseline interview and agreed to be contacted again. Interviews were carried out a year after the baseline interviews in waves that corresponded to the interview waves at

³ It should be noted that although the pilots were not operating in these areas the comparison sample could still have paid to attend childcare. As such the comparison being made in the evaluation was between children who took up a pilot place and children whose experiences represented the counterfactual of what would have happened in the absence of the pilot.

baseline. All the interviews were face-to-face and were supplemented by a short paper self-completion element for the parent and two cognitive assessments with the child.

As discussed in Section 1.1, the children in the pilot were often purposively selected because they were disadvantaged in some way, whereas the children in the comparison group were more broadly selected from disadvantaged areas. As such, although both groups of children were drawn from disadvantaged populations, the children in the comparison group were relatively less disadvantaged than the children in the pilot.

This needed to be addressed at the analysis stage because it was important to ensure that the pilot and comparison groups were as similar as possible when comparing outcomes. For this reason, the original impact assessment used a wide range of baseline variables to generate a propensity score on which to match the pilot group to the comparison group. After matching, any differences in outcomes between the two samples at age three were taken to represent the impact of the pilot.

The baseline and follow-up stages both collected a wide range of information including:

- Socio-demographic information
- Perceptions of the initiative from parents of children in the pilot
- Use of other formal and informal childcare
- Home learning environment
- Child-parent relationship
- Children's cognitive ability
 - Vocabulary⁴
 - Non-verbal reasoning⁵
- Children's social skills⁶

The variables and the matching technique are equally relevant within this research study and details of our analysis approach can be found in Section 1.4.

1.2.2 National Pupil Database

The primary outcome data for this research study was taken from the following elements of the National Pupil Database:

- a) The Early Years Census and concurrent Schools Census
- b) The Early Years Foundation Stage (EYFS) profile

⁴ As measured by the British Ability Scales (Elliot, Smith, and McCulloch 1996) and the Sure Start Language Measure (Roy et al. 2005).

⁵ As measured by the British Ability Scales (Elliot, Smith, and McCulloch 1996).

⁶ As measured by the Adaptive Social Behaviour Inventory (Hogan, Scott and Bauer 1992).

The Early Years and Schools Census both collect data on children using funded childcare provision during each academic year which means that they can be used to look at take-up of early education prior to starting school.⁷ This provides an indirect measure of the pilot's benefit because research has shown that attending high quality early education has a positive effect on children's outcomes (Sylva et al. 2004).

The Early Years and Schools Census data is collected in the spring of each academic year. Hence, the measure of take-up in this study shows whether children were receiving any early education in the spring of 2009 when they would have been aged either three or four (depending on their date of birth).⁸

The Early Years Foundation Stage (EYFS) was introduced in September 2008. The EYFS profile is an assessment based upon observations of children's learning and development. These must be completed for each child during the academic year when they reach the age of five (which is typically the Reception Year in primary school). The profile describes the child's level of attainment at the end of the EYFS and identifies their learning needs for the next stage of school to help Year 1 teachers plan effectively and appropriately. It includes six areas of learning made up of 13 assessments scales. Nine points can be achieved on each of these scales and so the total EYFS score ranges from 13 to 117.

When the children from this study were assessed the EYFS profile scored children across the following areas of development:⁹

- Personal, Social and Emotional Development
 - Dispositions and Attitudes
 - Social Development
 - Emotional Development
- Communication, Language and Literacy
 - Language for Communication and Thinking
 - Linking Sounds and Letters
 - Reading
 - Writing
- Problem Solving, Reasoning and Numeracy
 - Numbers as Labels and for Counting
 - Calculating
 - Shape, Space and Measures
- Knowledge and Understanding of the World

⁷ The Early Years Census is completed by the private, voluntary and independent sector. The Schools Census is completed by the maintained sector.

⁸ We did not look at the 2008 spring census because at that time a number of children in the sample were not yet eligible for early education for three and four year olds. Similarly we did not look at the 2010 spring census because at that time most children in the sample would already have started full-time school.

⁹ A simpler version of the EYFS was introduced in September 2012.

- Physical Development
- Creative Development.

The EYFS profile outcomes correspond reasonably well to those measured in the original evaluation. For example, the Personal, Social and Emotional Development scale in the EYFS profile corresponds well to the ASBI (which measures: pro-social behaviour; conformity/ compliance; confidence; anti-social behaviour; and anxiety). Similarly, the Language for Communication and Thinking subscale in the EYFS profile involves similar abilities to those required by the BAS-II naming vocabulary task and the SSLM (i.e. listening, responding to and using language). Finally, the Shape, Space and Measures, and Knowledge and Understanding of the World subscales in the EYFS profile involve similar abilities to those required by the BAS-II picture similarity task (i.e. recognising shapes, making comparisons, and understanding how everyday objects are used). In contrast, there are no obvious parallels between the Physical Development and Creative Development subscales of the EYFS profile and the constructs measured in the original evaluation.

The similarity between the outcomes measured at age three and the outcomes available at age five suggest that if the effects of the pilot have persisted then the follow-up analysis has a good chance of identifying them. This assumes that the EYFS assessments are sufficiently sensitive to identify the differences between children of different abilities and that the size of the effect is large enough to identify with the sample size available.¹⁰

1.3 Linking the original survey data to the NPD

The original evaluation included 2,372 children in the impact analysis. At the follow-up stage their parents were asked for their consent to link their survey data to administrative data held by the Department for Children, Schools and Families (now the Department for Education). In total 93% of parents gave their consent for data linkage meaning that 2,208 children were eligible for inclusion in this analysis of their NPD data.

The main challenge when attempting to undertake the data linkage was that the original survey data and the NPD contain no unique identifier common to both datasets. Because of this the two files had to be linked using a range of child characteristics. These were date of birth, forename, surname, postcode and gender. All these were available in the NPD and the original survey data.

The data linkage was carried out in a number of steps with a different combination of identifying variables used at each step in order to maximise the chances of finding a

¹⁰ With the available sample sizes for the overall analysis we could detect differences in percentages of around 6.2 percentage points when the prevalence of the outcome is 50%. We could detect differences in the mean of approximately 2.2 when the mean of the outcome (such as the EYFS score) is 83.2; or we could detect changes in the mean of 0.2 when the mean of the outcome (such as one of the EYFS learning areas) is 6.4. This assumes 80% power, and that all tests are two-sided and use a 5% significance level.

match. The first step involved using forename, surname and date of birth. As expected this combination uniquely identified most of the children on the NPD but it did not produce a match for every case.

Subsequent steps involved using different combinations of these variables. Later steps involved using a database of common nicknames and short versions of names (for example, to highlight a potential match between a child called Elizabeth in one file and Lizzie in another).

In some cases these combinations gave no match or multiple matches within the NPD. For that reason the later stage of data linkage also used gender to narrow the match and looked to see whether their postcode could identify a match (experience from longitudinal surveys suggests that families with young children don't tend to move very far).

In the end, 92% of children for whom we had consent for data linkage were successfully identified in the NPD. This meant that we had NPD data for a total of 2,039 children (1,080 comparison children and 959 children in the pilot). This represents 86% of the children from the original impact analysis.

1.4 Approach to analysis

1.4.1 Matching the pilot and comparison children

The analysis methods used in this study were chosen to be comparable to those used in the original evaluation. This was to make sure that any differences found were genuine and not due to changes in methodology.

In the original evaluation the pilot and comparison samples were matched using propensity score matching. This matching exercise produced a set of weights which, when applied to the data, made the profile of the baseline comparison sample match that of the baseline pilot sample. This method allows for multiple variables to be matched concurrently. Essentially, the difference between the two samples was modelled using logistic regression modelling (with all relevant baseline characteristics being predictors) and the modelled probability (or propensity) of being in the pilot sample was recorded per person. Children in the pilot were then matched to the comparison children in such a way that the two matched samples had equivalent propensity score profiles. This matching can be done in a number of ways. The original evaluation used 'kernel matching' whereby each pilot child was matched to a weighted distribution of comparison children, the weighting per comparison child being determined by the difference between their propensity score and the pilot child's propensity score.

The original matching weights could not be used in this research project because, as discussed in Section 1.3, we were only able to link NPD data to 86% of the families included in the original analysis. This means that 14% of the families that were included in the original evaluation were excluded from this study because of a lack of outcome data. The exclusion of these families would affect the ability of the weights to make the

pilot and comparison samples comparable (since the characteristics of the unmatched children could vary across the two groups).

For this reason we propensity score matched all families for whom we had outcome data available, using the same set of baseline variables used in the original evaluation.¹¹ As in the original evaluation, variables relating to the cognitive and social ability of the children were all entered into the logistic regression model. These included:

- Size of English vocabulary
- Size of vocabulary in another language
- Parental concern about child's language development and how child is learning
- Whether child had started to put words together yet
- Relative score on pro-social behaviour
- Relative score on anti-social behaviour
- Relative score on confidence
- Relative score on compliance
- Relative score on anxiety

All other variables were added to the model stepwise. This approach involves adding the variables to the model sequentially - starting with the variable that is most correlated with the outcome. The final model only retains variables that are related in a statistically significant way to the dependent variable. The variables that were added into the logistic regression model stepwise included:¹²

- Use of formal and informal childcare from age zero to age one
- Use of formal and informal childcare from age one to age two
- Intensity of the home learning environment
- Regularity of bed times
- Regularity of meal times
- Regularity of family meals
- Degree of TV watching
- Activity on child's birthday
- Frequency of visits to or by friends with children
- Frequency of attendance at parent and child groups
- Number of children in the household
- Family type (couple/ lone parent)
- Family work status
- Maternal work status
- Family socio-economic status
- Respondent's qualification level

¹¹ The only difference in the nature of the variables included in the propensity score matching was that the area level variables included in the original evaluation were excluded here because in this analysis they proved too powerful in the model and prevented our ability to match the samples on other characteristics which were more closely associated with the outcomes of interest.

¹² The final model for the overall sample can be found in 0.

- Teenage parenthood
- Whether either parent has a longstanding illness or disability
- Age of child at baseline
- Age of child within academic year
- Whether child speaks English as an additional language
- Whether child has special educational needs
- Whether child has a longstanding illness or disability
- Whether someone in the family receives Job Seekers Allowance
- Whether someone in the family receives Income Support
- Whether someone in the family receives housing benefit/ council tax benefit
- Whether someone in the family receives sickness and disability benefit
- Household income
- Housing tenure

A separate match was undertaken for each subgroup considered in Section 2.2. The quality of the overall match and the match for each subgroup was explored before undertaking the outcome analysis. In each case it appeared to achieve a good balance between the pilot and comparison samples across the baseline characteristics. (See Appendix C for illustrations of the effect that the matching had on the outcomes for the comparison group.)

Because of the nature of the study, we had little information about what happened to these children between the end of the pilot (aged three) and their EYFS profile assessment (age five). The data available included information on the type of childcare they had received at age three or four, quality of this provision as measured by Ofsted inspection grades, and information on the school they attended at the time of their EYFS assessment. This information was not included in the propensity matching because any experiences children had after the pilot could have been shaped by the pilot itself. We have, however, considered whether any differences in these experiences could explain the analysis findings in section 2.3.

1.4.2 Age at assessment

Age is strongly associated with children's cognitive ability. In particular, children's age at the point of the EYFS assessment is a strong determinant of their score (Crawford et al. 2007; Crawford et al. 2011). For this reason it is important to control for age when matching the pilot and comparison groups. As discussed in section 1.2.1 the original baseline survey data was collected between January 2007 and December 2007 and the follow-up survey data was collected between January 2008 and November 2008. The cognitive assessments at follow-up were timed to coincide approximately with children's third birthdays and so there was limited variation in the children's ages at follow-up.

In contrast, the Early Years Foundation Stage profile is completed at a fixed point in the summer term. As such, the children in the evaluation will have been different ages when

assessed (with those born in September 2004 being the oldest and those born in August 2005 being the youngest). Unfortunately this meant that there were substantial differences between the age distribution of the pilot and comparison groups as shown in the table below.

Table 1.1 Profile of children's month of birth			
<i>Base: All children</i>			
	Pilot children	Comparison children	Total
September 04	113	0	113
October 04	115	0	115
November 04	87	0	87
December 04	77	126	203
January 05	210	131	341
February 05	166	104	270
March 05	191	118	309
April 05	0	118	118
May 05	0	122	122
June 05	0	113	113
July 05	0	116	116
August 05	0	132	132
<i>Unweighted base</i>	959	1,080	2,039

Some children in the pilot were born between September and November 2004 but there are no equivalents in the comparison group. Some comparison children were born between April and August 2005 but there are no equivalents in the pilot group. These differences in children's ages at the point of the EYFS assessments were too stark to control for with analysis techniques and it was only possible to achieve a good match where the ages of children overlap.

This meant that the main analysis needed to exclude the very youngest and oldest children for whom we had no equivalents to match against. However, it was also important to strike a balance between excluding children whom we could not match well and maximising the sample size in order to maintain analytical power. In the end we decided to focus the analysis on children born between November 2004 and April 2005 because although there were no exact equivalents for children in the pilot born in November or for comparison children born in April we could merge November and December into one age category and do the same for March and April. This left 597 children in the comparison group and 731 in the pilot group.

The following section describes the impact findings for this group of children.

2 Impact at age five

2.1 Overall impact

Table 2.1 shows that the children in the pilot scored 81.5 on the EYFS profile and the matched comparison group scored 83.3. There is no statistically significant difference between these scores.¹³ The only statistically significant difference between the pilot and comparison groups in these six areas of learning was that the comparison group scored slightly higher on Knowledge and Understanding of the World (6.5 compared with 6.3).

There was no difference between the pilot and comparison samples in their take-up of early education (both 96%).

Table 2.1 Impact of the pilot on EYFS profile scores and take-up of early education			
<i>Base: All children born between Nov 04 and Apr 05</i>			
	Pilot children	Matched comparison children	p-value for difference ¹⁴
EYFS profile total score	81.5	83.3	0.266
- Personal, Social & Emotional Development	19.5	19.9	0.362
- Communication, Language & Literacy	23.9	24.4	0.315
- Problem Solving, Reasoning & Numeracy	18.9	19.4	0.169
- Knowledge & Understanding of the World	6.3	6.5	0.042
- Physical Development	6.8	6.8	0.657
- Creative Development	6.2	6.3	0.782
% Take-up of Early Education at Age 3 or 4	96%	96%	0.763
<i>Unweighted base</i>	731	597	
<i>Weighted base</i>	N/A	731	

These findings provide no evidence that the two year old pilot was effective in improving children's outcomes at age five, as measured by the EYFS profile, or their take-up of early education.

¹³ The EYFS assessment has six areas of learning that are comprised of 13 assessments scales. Nine points can be achieved on each of these scales and so the total EYFS profile score can range from 13 to 117. For more information see section 1.2.2.

¹⁴ The p value is for a linear regression test (comparing the pilot and comparison children).

2.2 Impact on subgroups

2.2.1 Children who attended better quality settings as part of the pilot

In the original evaluation, the overall lack of impact disguised the fact that for those children who were in better quality pilot settings there was an impact on their vocabulary and their relationship with their parents (Smith et al. 2009).¹⁵ This section considers whether these effects were still there when the children were aged five.

Of the 731 children in the pilot who were included in the overall analysis for this study:

- 45 attended a setting where its quality scored 5 or more on the Infant-Toddler Environment Ratings Scale (high quality);
- 103 attended a setting where its quality scored 4 (adequate quality);
- 75 attended a setting that scored 3 or less (low quality); and
- No information on quality of setting was available for 508 children.¹⁶

Table 2.2 replicates the analysis from the original evaluation. It compares the 148 children in the pilot who attended an average or high quality setting (ITERS score of 4 or more) with the whole matched comparison group.¹⁷ This shows that the children in the pilot who attended an average or high quality setting and the matched comparison sample achieved very similar scores across the EYFS profile as a whole and across the six areas of learning. None of the apparent differences were statistically significant.

There was no statistically significant difference in take-up of early education.

¹⁵ The University of Oxford used the Infant-Toddler Environment Rating Scale (ITERS) to conduct quality assessments in a sample of the pilot settings. These assessments were used to compare children in the pilot who went to a higher quality setting (i.e. settings that had scored 4 or more) with the comparison sample (see Smith et al. 2009 for more details).

¹⁶ This is because of the evaluation design – quality assessments were carried out with only a subsample of settings attended by children in the pilot group (Smith et al. 2009). It was not deemed appropriate to use imputed values for setting quality for the analysis of the pilot's effectiveness (either in the original evaluation or in this follow-up study).

¹⁷ As discussed in Chapter 1, the comparison group could have attended no childcare or childcare of varying quality.

Table 2.2 Impact of the pilot on EYFS profile scores and take-up of early education for children who went to better quality settings

Base: All children born between Nov 04 and Apr 05, who were either in the comparison group or in the pilot group and attended a setting rated 4+ on the ITERS scale

	Pilot children	Matched comparison children	p-value for difference
EYFS profile total score	83.2	83.1	0.959
- Personal, Social & Emotional Development	19.6	19.6	0.894
- Communication, Language & Literacy	24.5	24.5	0.991
- Problem Solving, Reasoning & Numeracy	19.4	19.4	0.911
- Knowledge & Understanding of the World	6.4	6.5	0.751
- Physical Development	6.9	6.8	0.749
- Creative Development	6.4	6.3	0.585
% Take-up of Early Education at Age 3 or 4	97%	95%	0.283
<i>Unweighted base</i>	<i>148</i>	<i>597</i>	
<i>Weighted base</i>	<i>N/A</i>	<i>148</i>	

The number of children in the pilot that we know attended a high quality setting (ITERS score of 5 or more) was very low (45), so this makes the estimates for this group less precise. Nevertheless, we tried comparing the group of children who attended a high quality setting with the whole matched comparison group. The findings from this analysis also found no significant differences.

Trends by quality of setting within the pilot group

An alternative way of looking at the effect of attending different quality settings is to see whether there are any trends in children's EYFS profile scores that are associated with the quality of the setting they attended.

In Table 2.3, we compared the outcomes for children in the pilot who attended a high quality setting (ITERS score of 5 or more) with children in the pilot who attended an adequate or low quality setting (ITERS score of 4 or less) (after matching the profiles of the two groups of children in order to make the comparison valid). As discussed above, the number of children that we know attended high quality settings was very low (45). This makes the estimates for this group less precise and means that a larger difference between the groups is needed to register as statistically significant. However, Table 2.3 shows a consistent trend whereby the children in the pilot who attended a high quality setting scored better on the EYFS assessment than children in the pilot who attended an adequate or low quality setting.

In particular we can see that the children in the pilot who attended a high quality setting scored higher on the Communication, Language and Literacy subscale (25.8 compared with 24.0) and on the Creative Development subscale (6.6 compared with 6.2). This positive finding for the Communication, Language and Literacy subscale is consistent with the findings from the original evaluation which showed that children who attended a better quality setting had a larger vocabulary than children in the comparison group (Smith et al. 2009).

Table 2.3 Variation across EYFS profile scores and take-up of early education for pilot children attending settings of different quality			
<i>Base: All pilot children born between Nov 04 and Apr 05, who attended a setting whose quality was measured on the ITERS scale</i>			
	Low/ adequate (<=4)	High (>=5)	p-value for difference
EYFS profile total score	82.4	86.0	0.137
- Personal, Social & Emotional Development	19.7	20.2	0.338
- Communication, Language & Literacy	24.0	25.8	0.049
- Problem Solving, Reasoning & Numeracy	19.3	19.8	0.396
- Knowledge & Understanding of the World	6.3	6.4	0.668
- Physical Development	6.8	7.0	0.328
- Creative Development	6.2	6.6	0.030
% Take-up of Early Education at Age 3 or 4	96%	98%	0.575
<i>Unweighted base</i>	<i>178</i>	<i>45</i>	
<i>Weighted base</i>	<i>45</i>	<i>N/A</i>	

We also compared the group of children who attended a low quality setting (ITERS score of 3 or less) with those who attended an adequate or high quality setting (ITERS score of 4 or more). However the trend seen in Table 2.3, where children in the high quality settings scored better, was not evident when we grouped the quality scores in this way.

This analysis provides some evidence that *high* quality settings (those with an ITERS score of 5 or more) made a difference to children who attended the pilot. However the limited number of children who attended a high quality setting makes it difficult to pick up an impact on this group when comparing the pilot and matched comparison group.

2.2.2 Children from black or minority ethnic backgrounds

This section considers whether the pilot had an impact specifically on children from a black or minority ethnic (BME) background.

In total, only 151 children in the comparison group and 149 children in the pilot were from a BME background. Unfortunately, this means that although we would have liked to look

at different minority groups separately, it was only possible to look at BME groups as a whole.

Table 2.4 shows that the scores on the EYFS profile overall and across the six learning areas were very similar for BME children in the pilot and those in the comparison group.

However, while 98% of BME children in the pilot took up early education at age three or four, this was the case for only 92% of BME children in the comparison group, giving a positive impact of six percentage points. Other research has shown that although take-up of early years education at ages three and four is very high, take-up is lower among children from a BME background (Speight et al. 2010). The finding here suggests that the pilot has helped improve take-up among children from a BME background.

Table 2.4 Impact of the pilot on EYFS profile scores and take-up of early education for children with a BME background			
<i>Base: All children born between Nov 04 and Apr 05, who have a BME background</i>			
	Pilot children	Matched comparison children	p-value for difference
EYFS profile total score	83.5	84.1	0.836
- Personal, Social & Emotional Development	19.9	20.0	0.829
- Communication, Language & Literacy	24.7	24.8	0.924
- Problem Solving, Reasoning & Numeracy	19.4	19.7	0.654
- Knowledge & Understanding of the World	6.3	6.4	0.760
- Physical Development	6.9	6.8	0.666
- Creative Development	6.3	6.3	0.761
% Take-up of Early Education at Age 3 or 4	98%	92%	0.046
<i>Unweighted base</i>	<i>149</i>	<i>151</i>	
<i>Weighted base</i>	<i>N/A</i>	<i>149</i>	

2.2.3 Children in other subgroups

The study also examined whether the pilot had an impact specifically on children in the following subgroups:

- Children who had lower cognitive ability at age two (identified as those who scored below the median on the Sure Start Language Measure at baseline (Roy et al. 2005)¹⁸
- Children who lived in one of the 20% most disadvantaged areas of the country at baseline (as measured by the Income Deprivation Affecting Children Index)¹⁹

¹⁸ The median score among the evaluation sample was 65 out of 100. In total, 285 children in the comparison group and 372 children in the pilot group scored below this threshold. These children formed the base of the analysis.

- Children who received more than 7.5 hours of early education as part of the pilot.²⁰

There were no statistically significant differences between the pilot sample and the comparison sample within these three subgroups on the EYFS profile either in the total score or the scores for the six learning areas. Similarly, there were no statistically significant differences in take-up of early education at age three or four (see tables in 0).

2.3 Explaining the results

2.3.1 Methodological issues

The methodologies of this follow-up study, and that of the original evaluation, were well suited for the purpose of examining the overall impact of the pilot on children who were offered places in early years settings. However, the design of the original evaluation (combined with the fact that we could not use all of the original sample for the follow-up study for reasons discussed in section 1.3) limited our ability to detect benefits of attending *high quality* provision as part of the pilot.

The original evaluation found that cognitive development scores were the highest for children who went to early years settings of the highest quality (ITERS score of 5 or more; see Smith et al. 2009, p. 104). In this study, very few children attended settings which had an ITERS score of 5 or more.²¹ Our analysis provides some evidence that children in the pilot who went to *high quality* settings achieved higher EYFS profile scores than children in the pilot who went to lower quality settings (see section 0). However, the limited number of children who attended a high quality setting makes it difficult to pick up an impact of the pilot when comparing children who attended high quality settings in the pilot with the matched comparison group.²²

Another important issue is whether the children in the pilot and those in the matched comparison group were well-matched at age five. Because of how children were recruited to the study,²³ the children in the pilot were relatively more disadvantaged than those in the comparison group. The propensity score matching used information from the

¹⁹ As might be expected given the targeting of the pilot, a substantial proportion of children from the evaluation sample lived in disadvantaged areas. Indeed, this was the case for 301 children in the comparison sample and 400 children in the pilot sample.

²⁰ As discussed in Chapter 1, the majority of children received 7.5 hours of early years education per week as part of the pilot, but some local authorities offered 9 or 12.5 hours. Indeed, of the 731 children in the pilot, 435 received exactly 7.5 hours and 296 received more hours.

²¹ Out of 731 children in the pilot group used for this study, we had information on the quality of settings for 223 children. This is because the evaluation was not designed to measure quality of *all* settings attended by children in the pilot (see Smith et al. 2009). Out of these 223 children, 45 attended settings of high quality (ITERS score of 5 or more).

²² The study has tried to match the 45 children in the pilot group who attended high quality settings with children in the comparison group. It was not possible to achieve a good match given the small sample size and the level of differences between the pilot group and the comparison group.

²³ The pilot group was selected from children who were offered pilot places by local authorities. Local authorities based their selection on different indicators of disadvantage. In contrast, the comparison sample was selected from families living in relatively deprived areas where the pilot was not operating and which had a relatively large minority ethnic population. (For more detail about the evaluation design, see Smith et al. 2009.)

baseline survey (that is, when the children in the study were aged two). Our tests showed that we were successful in achieving a good statistical match between the two groups *at baseline*. However, there may have been unobserved differences between the two groups that could not be controlled for. Also, the children in the study could have had very different experiences since they were two, and these experiences might have affected their attainment in their EYFS profile assessments.

Because of the nature of the study, we had little information about what happened to the children between the end of the pilot at age three and their EYFS assessment at age five. Moreover, even if more information about the intervening period were available, it would not be valid to use it when matching the pilot and comparison groups as the children’s experiences might have been affected by the pilot itself. Instead, we explored whether the experiences of the pilot and comparison groups between the ages of three and five were different to understand whether this might help to explain the findings. We were able to look at two areas:

- The types of early years settings attended by the children when they were aged three or four, and the quality of that provision, as measured by Ofsted inspections; and
- The schools they attended for their Reception Year.

2.3.2 Early years provision received at age three to four

There was no statistically significant difference in the type of early education setting attended by children in the pilot and comparison group when they were aged three or four years old (see Table 2.5). However, the trend was towards children in the pilot being somewhat more likely to attend a maintained setting (e.g. a maintained school setting such as a nursery class attached to a primary school, or a local authority day nursery) and less likely to attend a private, voluntary or independent setting.

Table 2.5 Type of early education setting attended at age three/four ²⁴		
<i>Base: All children born between Nov 04 and Apr 05</i>		
	Pilot children	Matched comparison children
	%	%
Maintained school setting (e.g. maintained nursery school or nursery class attached to a primary school)	67	64
Private early years setting	18	20
Voluntary / independent / other early years setting	7	11
Local authority day nursery	4	1
No childcare	4	4

²⁴ Eighteen children in the study attended both maintained school settings and early years settings for their early years provision. They are shown in the table as attending maintained school settings.

Table 2.5 Type of early education setting attended at age three/four ²⁴		
p-value ²⁵		0.111
<i>Unweighted base</i>	731	597
<i>Weighted base</i>	n/a	731

Evidence from other studies suggests that the quality of early years education in maintained settings tends to be better than that in private or voluntary settings, while the quality in voluntary settings tends to be similar to that in private settings (e.g. Mathers and Sylva 2007; Mathers et al. 2007).

The direction of the trend found suggests that, on average, the quality of provision received by children in the pilot when they were aged three to four might have been somewhat better than it was among the comparison group. To explore this hypothesis further, the study compared settings' Ofsted inspection grades.²⁶

Table 2.6 shows distribution of settings' Ofsted grades among the pilot and comparison groups of children. The results suggest that the early years provision received by children in the pilot group when they were aged three to four was of somewhat better quality than that received by children in the comparison group (the difference is statistically significant at the 5% level). For example, 18% of children in the pilot group went to settings which received an 'outstanding' grade from Ofsted, and 65% went to those which were considered 'good', while the respective figures for the comparison group were 14% and 63%.

²⁵ The p value is for a chi-squared test (comparing the pilot and comparison children).

²⁶ NatCen Social Research was supplied with Ofsted data for maintained school settings and early years settings by DfE. For maintained school settings, the data covered academic years from 2005/06 to 2011/12. For early years settings, the data covered the period from January 2006 to September 2010. As settings are not inspected every year, the analysis used an Ofsted grade which was nearest in time to May 2009 (as our measurement of take-up of early years provision at age three/four was from Spring 2009 census). For maintained school settings, we used a grade for the school's overall effectiveness of the EYFS (which would have taken into account provision for pupils in nursery schools/classes and in the Reception Year). For early years settings, we used a grade for their overall effectiveness. It is worth noting that a change in how early years settings were inspected by Ofsted took place on 1 September 2008 with the introduction of the EYFS. Prior to this date, there were three types of inspections: one for 'care', one for 'nursery education', and 'integrated inspections' for providers that provided both care and nursery education (so both were assessed). From 1 September 2008, providers received an Early Years Register inspection, which took account of the requirements of the EYFS framework. The variety (1) in the types of settings attended by the children, (2) in the types of Ofsted inspections providing the grades and (3) in how contemporaneous Ofsted inspection grades were to when the children actually were at the setting should be taken into account when interpreting the findings, as they limit comparability among Ofsted grades used in the analysis.

Table 2.6 Ofsted inspection grades for settings attended by children at age three/four ²⁷		
<i>Base: Children born between Nov 04 and Apr 05 who received early years provision at age three/four</i>		
	Pilot children	Matched comparison children
	%	%
Outstanding	18	14
Good	65	63
Satisfactory	17	19
Inadequate	+	4
p-value ²⁸		0.043
<i>Unweighted base</i>	675	533
<i>Weighted base</i>	n/a	669

These results suggest that the experience of early years provision which the children in the study had when they were aged three to four probably does not explain why there was no evidence of an improvement in the EYFS profile scores of the pilot group at age five.

2.3.3 School experience during the Reception Year

Children's learning at school during the Reception Year and the effectiveness of schools have direct relevance to how well the children perform in their EYFS profile assessments at the end of that year. The study attempted to explore whether there was any evidence of differences in the characteristics and performance of schools attended by children in the pilot group compared with schools attended by children in the comparison group.

Table 2.7 shows various characteristics of schools attended by children in the study for their Reception Year. Namely, the study examined schools' socio-demographic composition,²⁹ schools' results on the EYFSP³⁰ and on Key Stage 1 (KS1),³¹ as well as schools' value added scores (from EYFS to KS1).³²

²⁷ Those children who attended both maintained school settings and early years settings for their early years provision (n=18) were classified as attending maintained settings, and we used Ofsted grades for their maintained settings in the analysis of quality of the provision they received.

²⁸ The p value is for a chi-squared test (comparing the pilot and comparison children).

²⁹ Schools' socio-demographic composition data was extracted from Edubase.

³⁰ Schools' mean EYFS profile total scores were calculated from pupil level data for academic year 2009/10, which was supplied to NatCen Social Research by DfE.

³¹ KS1 data was downloaded from DfE performance tables available online.

³² School level mean value added scores were calculated by DfE using children's EYFS total pointscore across the profile and their KS1 total pointscore in reading, writing and maths. This was based on all children who were assessed on the EYFS in 2008 and all children who were assessed at KS1 in 2010 at each school that was attended by the samples of pilot and comparison children in the study.

Table 2.7 Characteristics of school attended at age five			
<i>Base: All children born between Nov 04 and Apr 05</i>			
	Pilot children	Matched comparison children	p value ³³
Socio-demographic composition			
% of pupils with English as an additional language	18%	26%	0.017
% pupils eligible for free school meals	29%	29%	0.982
Pupils' academic achievements (school average)			
Mean EYFSP total score	82.0	82.9	0.365
Mean KS1 points score	14.1	14.4	0.104
% of cohort below expected level at end of KS1	27%	24%	0.062
% of cohort above expected level at end of KS1	17%	19%	0.025
Mean Value Added score (from EYFS to KS1) ³⁴	0.040	-0.002	0.004
<i>Unweighted base for % of pupils with EAL</i>			
	620	541	
<i>Weighted base for % of pupils with EAL</i>			
	N/A	659	
<i>Unweighted base for % of pupils eligible for FSM</i>			
	709	582	
<i>Weighted base for % of pupils eligible for FSM</i>			
	N/A	716	
<i>Unweighted base for EYFSP total score</i>			
	731	597	
<i>Weighted base for EYFSP total score</i>			
	N/A	731	
<i>Unweighted base for KS1 measures</i>			
	593	492	
<i>Weighted base for KS1 measures</i>			
	N/A	609	
<i>Unweighted base for Value Added score</i>			
	720	593	
<i>Weighted base for Value Added score</i>			
	N/A	716	

With regard to schools' socio-demographic composition, there were no differences between the pilot and the comparison group in the percentage of pupils at their schools who were eligible for free school meals (29% for both; see Table 2.7) However, schools attended by the pilot group had a significantly lower percentage of pupils for whom English was not their first language (18%, compared with 26% among schools attended by the comparison group).³⁵ As such, it could be suggested that schools attended by the comparison group were working with a more challenging population of children than those attended by the pilot group, which is useful context for interpreting schools' attainment results.

³³ The p value is for a linear regression test (comparing the pilot and comparison children).

³⁴ Positive values at the value added score mean that when prior attainment (FSP score) was controlled for, schools were adding more value to pupils' KS1 scores than the national average; and negative values at the score mean that schools were adding less value than the national average.

³⁵ This difference in the figures may be linked to the original design of the evaluation – the comparison sample was selected from families living in areas which had a relatively large minority ethnic population (Smith et al. 2009).

Turning now to schools' attainment figures, they suggest that mean EYFSP total scores did not differ significantly between schools attended by the pilot and the comparison group (82.0 and 82.9 points respectively).³⁶ However, schools attended by the pilot group had somewhat lower KS1 scores than those attended by the comparison group. For example, in schools attended by children in the pilot, marginally more pupils fell below the expected level for KS1 than in comparison children's schools (27% compared with 24%). Likewise, in schools attended by children in the pilot, fewer pupils exceeded the expected level for Key Stage 1 than in comparison children's schools (17% compared with 19%). This suggests that perhaps schools attended by the pilot group were somewhat less well performing than those attended by the comparison group (at least, at KS1), especially if we take into account that schools attended by the comparison group were dealing with a more challenging group of pupils and still managed to achieve somewhat better KS1 results.

This conclusion, however, is not consistent with the analysis of schools' value added scores. The mean value added scores (from EYFS to KS1) were 0.040 for schools attended by the pilot group and -0.002 for schools attended by the comparison group. The difference was small in absolute terms but it was statistically significant at the 1% level and suggests that schools attended by the pilot group were adding more value to their pupils' KS1 results once prior attainment (EYFS profile score) was controlled for.

As the evidence is somewhat contradictory, we cannot conclude that children's school experience during the Reception Year was systematically different between the pilot and the comparison group and thus there is no evidence that it contributed to the apparent erosion of the pilot's positive impact by age five.

³⁶ We also compared mean scores for the six subscales of the EYFSP, and found no statistically significant differences there either (table not shown).

3. Conclusions

At the age of five, overall there were no statistically significant differences between the children who had taken part in the pilot and the children in the comparison sample in their total EYFS profile score (81.5 and 83.3 respectively).

The same was true for:

- Children from a black or minority ethnic background;
- Children who had lower than average cognitive ability at age two;
- Children who received more than 7.5 hours of early education as part of the pilot; and
- Children who lived in one of the 20% most disadvantaged areas of the country at age two.

The **original evaluation** identified the importance of providing pilot places in better quality settings (Smith et al. 2009), so we undertook some comparable subgroup analysis that looked specifically at the children in the pilot who went to better quality settings. This analysis found no differences between the children in the pilot who went to better quality settings (ITERS score of 4 or more) and the matched comparison sample.

There were, however, some methodological issues, which affected our ability to detect the effect of the pilot on the subgroup of children who attended better quality settings. The original evaluation found that cognitive development scores were the highest for children who went to early years settings of the highest quality (ITERS score of 5 or more; see Smith et al. 2009, p. 104). In this study, very few children attended settings which had ITERS scores of 5 or more. This limited the study's ability to pick up the impact of the pilot on the subgroup of children who went to better quality settings compared with children in the matched comparison group.

Further analysis comparing children **within the pilot group** who attended high quality settings (ITERS score of 5 or more) with those who attended adequate or low quality settings (ITERS score of 4 or less) found that the high quality group scored significantly higher across the Communication, Language and Literacy subscale (25.8 compared with 24.0) and also across the Creative Development subscale (6.6 compared with 6.2). This provides some evidence that *high* quality settings made a positive difference to children who attended the pilot. However, the small number of children in the sample who attended a high quality setting makes it difficult to pick up this effect when comparing the pilot and matched comparison group.

In terms of children's take-up of early education at age three or four, there were no differences between the pilot and comparison children overall. Similarly, there were no differences across most of the subgroups. The exception was for children from a black or minority ethnic background where we found a positive impact of six percentage points on take-up of early education (98% of BME children in the pilot had taken up early years

education compared with 92% of BME children in the matched comparison sample). Other research has shown that although take-up of early years education at ages three and four is very high, take-up is lower among children from a BME background (Speight et al. 2010). The finding here suggests that the pilot has helped improve take-up among children from a BME background.

The study also considered whether the children in the pilot and those in the matched comparison group were still well-matched at age five, or whether there were systematic differences in the experiences they had had by then.³⁷ We were able to look at (1) the type and quality of early years settings (as measured by Ofsted inspection grades) attended by the children when they were aged three or four; and (2) the schools they attended for their Reception Year.

Our exploratory analysis provided no evidence of significant differences in the types of early years settings that the pilot and comparison groups attended. However, there was some evidence that the children in the pilot group went to settings of better quality than those in the comparison group. Analysis of school performance data was inconclusive and did not provide evidence of systematic differences in the quality of schools attended by the two groups during the Reception Year. To sum up, these exploratory findings do not help us explain why there was no evidence of an improvement in the EYFS profile scores of the pilot group at age five when compared with children in the matched comparison group.

³⁷ The pilot and comparison groups of children were matched according to baseline characteristics at age two.

References

Crawford, C., Dearden, L., and Meghir, C. (2007) *When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England*. IFS. London.

Crawford, C., Dearden, L., and Greaves, E. (2011) *Does when you are born matter? The impact of month of birth on children's cognitive and non-cognitive skills in England*. IFS Briefing Note BN122. IFS. London.

Dickens S., Wollny, I., and Ireland, E. (2012) *Childcare Sufficiency and Sustainability in Disadvantaged Areas*. DfE Research Report DFE-RR246.

Elliot, C. D., Smith, P., and McCulloch, K. (1996). *British ability scales: Second edition (BAS II)*. Windsor, UK: nferNelson.

Gibb, J., Jelcic, H., La Valle, I., Gowland, S., Kinsella, R., Jessiman, P., Ormston, R. (2011) *Rolling out free early education for disadvantaged two year olds: an implementation study for local authorities and providers*. Department for Education, Research Report DfE-RR131.

HM Treasury (2004) *Choice for Parents, the Best Start for Children: a Ten Year Strategy for Childcare*, London: The Stationery Office.

Hogan, A.E., Scott, K.G. & Bauer, C.R. (1992). 'The Adaptive Social Behavior Inventory (ASBI): A new assessment of social competence in high risk three year olds', *Journal of Psychoeducational Assessment*, 10, 230-239.

Mathers, S., and Sylva, K. (2007) *National evaluation of the Neighbourhood Nurseries Initiative: The relationship between quality and children's behavioural development*, DfES.

Mathers, S., Sylva, K, Joshi, H., Hansen, K., Plewis, I., Johnson, J., George, A., Linskey, F., and Grabbe, Y. (2007) *Quality of Childcare Settings in the Millennium Cohort Study*, Research Report SSU/2007/FR/025, DfES.

Melhuish, E. (2004) *A Literature Review of the Impact of Early Years Provision on Young Children, with emphasis given to Children from Disadvantaged Backgrounds*. Report for the Audit Commission. London: Birkbeck College, University of London.

Roy, P., Kersley, H., and Law, J. (2005) *The Sure Start Language Measure Standardisation Study*.

Smith, R., Purdon, S., Schneider, V., La Valle, I., Wollny, I., Owen, R., Bryson, C., Mathers, S., Sylva, K. and Lloyd E. (2009) *Early Education Pilot for Two Year Old Children Evaluation*. DCSF Research Report No. DCSF-RR134.

Speight, S., Smith, R., and Lloyd, E. with Coshall, C. (2010) *Towards universal early year's provision: analysis of take-up by disadvantaged families from recent annual childcare surveys*. DfE Research Report DfE-RR066.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B. (2004) *The effective provision of pre-school education (EPPE) project: final report: A longitudinal study funded by the DfES 1997-2004*, London: DfES/Institute of Education, University of London.

Appendix A Propensity score model for overall sample

Appendix table A.1 Propensity score model to model the likelihood of being in the pilot group			
<i>Base: All children born between Nov 04 and Apr 05</i>			
	Odds ratio	SE	p value
Child's English vocabulary (lowest decile)			
2nd decile	0.62	0.20	0.13
3rd decile	0.72	0.24	0.34
4th decile	0.47	0.17	0.04
5th decile	0.47	0.17	0.03
6th decile	0.40	0.14	0.01
7th decile	0.48	0.17	0.04
8th decile	0.60	0.23	0.18
9th decile	0.50	0.19	0.07
10th decile	0.76	0.30	0.48
Child's foreign vocabulary (zero)			
Below average	0.72	0.18	0.20
Above average	0.61	0.16	0.07
Parental concern about child's language development and how child is learning (none)			
One area of concern	1.49	0.31	0.05
More than one area of concern	1.43	0.43	0.24
Whether child is putting words together in sentences (not at all)			
Sometimes	0.74	0.23	0.33
Often	0.84	0.28	0.60
Child's pro-social behaviour (lowest decile)			
2nd decile	1.19	0.35	0.55
3rd decile	1.37	0.41	0.30
4th decile	0.77	0.24	0.40
5th decile	0.52	0.16	0.03
6th decile	0.82	0.24	0.50
7th decile	0.92	0.28	0.78
8th decile	0.63	0.20	0.14
9th decile	0.97	0.30	0.93
10th decile	0.54	0.17	0.05
Child's anti-social behaviour (lowest decile)			
2nd decile	0.75	0.22	0.32
3rd decile	0.86	0.25	0.59
4th decile	0.56	0.16	0.04
5th decile	0.92	0.28	0.78
6th decile	0.59	0.17	0.08
7th decile	0.66	0.20	0.16
8th decile	0.51	0.15	0.03

Appendix table A.1 Propensity score model to model the likelihood of being in the pilot group			
9th decile	0.72	0.22	0.27
10th decile	0.53	0.16	0.04
Child's compliance behaviour (lowest decile)			
2nd decile	1.31	0.38	0.36
3rd decile	0.70	0.21	0.24
4th decile	0.81	0.24	0.48
5th decile	0.86	0.26	0.63
6th decile	0.77	0.23	0.38
7th decile	1.05	0.32	0.87
8th decile	0.89	0.28	0.72
9th decile	0.69	0.21	0.24
10th decile	0.73	0.23	0.33
Child's anxiety behaviour (lowest decile)			
2nd decile	0.96	0.28	0.90
3rd decile	0.81	0.23	0.47
4th decile	0.90	0.26	0.71
5th decile	1.23	0.36	0.48
6th decile	0.94	0.27	0.82
7th decile	1.01	0.30	0.96
8th decile	1.18	0.35	0.58
9th decile	0.92	0.27	0.77
10th decile	1.48	0.44	0.19
Child's confidence behaviour (lowest decile)			
2nd decile	1.08	0.33	0.81
3rd decile	1.15	0.36	0.64
4th decile	1.11	0.35	0.74
5th decile	1.16	0.37	0.63
6th decile	1.52	0.49	0.20
7th decile	1.07	0.34	0.85
8th decile	1.54	0.50	0.18
9th decile	1.17	0.38	0.63
10th decile	1.86	0.62	0.06
Child's age in months at age two assessments (25 months)			
26 months	4.93	1.31	0.00
27 months	3.46	0.87	0.00
28+ months	15.93	4.39	0.00
Child's month of birth (Apr/Mar)			
February	3.54	0.74	0.00
January	1.58	0.28	0.01
November/ December	2.71	0.49	0.00
Household income (unknown)			

Appendix table A.1 Propensity score model to model the likelihood of being in the pilot group

£9,999 or less	1.47	0.43	0.19
£10,000 - £19,999	1.55	0.45	0.13
£20,000 - £29,999	1.49	0.45	0.19
£30,000+	0.58	0.18	0.08
Receipt of sickness and disability benefit (does not receive)			
Receives	2.31	0.54	0.00
Receipt of childcare between age one and age two (formal care only)			
Informal care only	1.92	0.40	0.00
Formal and informal care	1.24	0.29	0.36
No childcare	2.11	0.46	0.00

Note: Odd ratio >1 indicates higher odds of being in the pilot group, and odd ratio <1 indicates lower odds of being in the pilot group, compared to the reference category in bold and brackets.

Appendix B Additional tables for subgroup analysis

Appendix table B.1 Impact of the pilot on EYFS profile scores and take-up of early education for children with lower cognitive ability at age two

Base: All children born between Nov 04 and Apr 05, who had lower cognitive ability at age two

	Pilot children	Matched comparison children	p-value for difference
EYFS profile total score	75.3	78.6	0.212
- Personal, Social & Emotional Development	18.2	19.0	0.160
- Communication, Language & Literacy	21.7	22.6	0.309
- Problem Solving, Reasoning & Numeracy	17.5	18.2	0.290
- Knowledge & Understanding of the World	5.8	6.2	0.074
- Physical Development	6.3	6.6	0.150
- Creative Development	5.8	6.0	0.482
% Take-up of Early Education at Age 3 or 4	96%	95%	0.697
<i>Unweighted base</i>	372	285	
<i>Weighted base</i>	N/A	372	

Appendix table B.2 Impact of the pilot on EYFS profile scores and take-up of early education for children who lived in the 20% most disadvantaged areas of the country at age two

Base: All children born between Nov 04 and Apr 05, who lived in the 20% most disadvantaged areas of the country at age two

	Pilot children	Matched comparison children	p-value for difference
EYFS profile total score	80.2	82.6	0.312
- Personal, Social & Emotional Development	19.1	19.5	0.544
- Communication, Language & Literacy	23.4	24.4	0.191
- Problem Solving, Reasoning & Numeracy	18.6	19.3	0.168
- Knowledge & Understanding of the World	6.2	6.5	0.169
- Physical Development	6.8	6.8	0.960
- Creative Development	6.2	6.2	0.962
% Take-up of Early Education at Age 3 or 4	97%	96%	0.726
<i>Unweighted base</i>	400	301	
<i>Weighted base</i>	N/A	400	

Appendix table B.3 Impact of the pilot on EYFS profile scores and take-up of early education for children who received more than 7.5 hours of early education as part of the pilot

Base: All children born between Nov 04 and Apr 05, who were either in the comparison group or in the pilot group and received more than 7.5 hours of early education

	Pilot children	Matched comparison children	p-value for difference
EYFS profile total score	80.3	81.2	0.771
- Personal, Social & Emotional Development	19.2	19.5	0.557
- Communication, Language & Literacy	23.6	23.6	0.958
- Problem Solving, Reasoning & Numeracy	18.5	18.8	0.707
- Knowledge & Understanding of the World	6.2	6.3	0.723
- Physical Development	6.7	6.8	0.812
- Creative Development	6.2	6.2	0.936
% Take-up of Early Education at Age 3 or 4	96%	95%	0.701
<i>Unweighted base</i>	296	597	
<i>Weighted base</i>	N/A	296	

Appendix C Matched and unmatched EYFS profile scores and take-up of early education at age three or four

Appendix table C.1 Matched and unmatched EYFS profile scores and take-up of early education at age three or four			
<i>Base: All children born between Nov 04 and Apr 05</i>			
	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	81.5	83.3	86.7
- Personal, Social & Emotional Development	19.5	19.9	20.5
- Communication, Language & Literacy	23.9	24.4	25.8
- Problem Solving, Reasoning & Numeracy	18.9	19.4	20.2
- Knowledge & Understanding of the World	6.3	6.5	6.7
- Physical Development	6.8	6.8	7.1
- Creative Development	6.2	6.3	6.5
% Take-up of Early Education at Age 3 or 4	96%	96%	95%
<i>Unweighted base</i>	731	597	597
<i>Weighted base</i>	N/A	731	N/A

Appendix table C.2 Matched and unmatched EYFS profile scores and take-up of early education for children who went to better quality settings

Base: All children born between Nov 04 and Apr 05, who were either in the comparison group or in the pilot group and attended a setting rated 4+ on the ITERS scale

	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	83.2	83.1	86.7
- Personal, Social & Emotional Development	19.6	19.6	20.5
- Communication, Language & Literacy	24.5	24.5	25.8
- Problem Solving, Reasoning & Numeracy	19.4	19.4	20.2
- Knowledge & Understanding of the World	6.4	6.5	6.7
- Physical Development	6.9	6.8	7.1
- Creative Development	6.4	6.3	6.5
% Take-up of Early Education at Age 3 or 4	97%	95%	95%
<i>Unweighted base</i>	148	597	597
<i>Weighted base</i>	N/A	148	N/A

Appendix table C.3 Matched and unmatched EYFS profile scores and take-up of early education for children with lower cognitive ability at age two

Base: All children born between Nov 04 and Apr 05, who had low cognitive ability at age two

	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	75.3	78.6	82.6
- Personal, Social & Emotional Development	18.2	19.0	19.7
- Communication, Language & Literacy	21.7	22.6	24.1
- Problem Solving, Reasoning & Numeracy	17.5	18.2	19.2
- Knowledge & Understanding of the World	5.8	6.2	6.5
- Physical Development	6.3	6.6	6.9
- Creative Development	5.8	6.0	6.2
% Take-up of Early Education at Age 3 or 4	96%	95%	94%
<i>Unweighted base</i>	372	285	285
<i>Weighted base</i>	N/A	372	N/A

Appendix table C.4 Matched and unmatched EYFS profile scores and take-up of early education for children with a BME background

Base: All children born between Nov 04 and Apr 05, who have a BME background

	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	83.5	84.1	85.8
- Personal, Social & Emotional Development	19.9	20.0	20.3
- Communication, Language & Literacy	24.7	24.8	25.5
- Problem Solving, Reasoning & Numeracy	19.4	19.7	19.9
- Knowledge & Understanding of the World	6.3	6.4	6.6
- Physical Development	6.9	6.8	7.0
- Creative Development	6.3	6.3	6.4
% Take-up of Early Education at Age 3 or 4	98%	92%	93%
<i>Unweighted base</i>	149	151	151
<i>Weighted base</i>	N/A	149	N/A

Appendix table C.5 Matched and unmatched EYFS profile scores and take-up of early education for children who lived in the 20% most disadvantaged areas of the country at age two

Base: All children born between Nov 04 and Apr 05, who lived in the 20% most disadvantaged areas of the country at age two

	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	80.2	82.6	85.5
- Personal, Social & Emotional Development	19.1	19.5	20.2
- Communication, Language & Literacy	23.4	24.4	25.4
- Problem Solving, Reasoning & Numeracy	18.6	19.3	19.9
- Knowledge & Understanding of the World	6.2	6.5	6.7
- Physical Development	6.8	6.8	7.0
- Creative Development	6.2	6.2	6.4
% Take-up of Early Education at Age 3 or 4	97%	96%	95%
<i>Unweighted base</i>	400	301	301
<i>Weighted base</i>	N/A	400	N/A

Appendix table C.6 Matched and unmatched EYFS profile scores and take-up of early education for children who received more than 7.5 hours of early education as part of the pilot

Base: All children born between Nov 04 and Apr 05, who were either in the comparison group or in the pilot group and received more than 7.5 hours of early education

	Pilot children	Matched comparison children	Unmatched comparison children
EYFS profile total score	80.3	81.2	86.7
- Personal, Social & Emotional Development	19.2	19.5	20.5
- Communication, Language & Literacy	23.6	23.6	25.8
- Problem Solving, Reasoning & Numeracy	18.5	18.8	20.2
- Knowledge & Understanding of the World	6.2	6.3	6.7
- Physical Development	6.7	6.8	7.1
- Creative Development	6.2	6.2	6.5
% Take-up of Early Education at Age 3 or 4	96%	95%	95%
<i>Unweighted base</i>	296	597	597
<i>Weighted base</i>	N/A	296	N/A



Department
for Education

© NatCen Social Research [March 2013]

Ref: DFE- RR225

ISBN: 978-1-78105-218-1

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Any enquiries regarding this publication should be sent to us at Adele Rowe, Ground Floor, Sanctuary Buildings, Great Smith Street, London, SW1P 3BT

Email: adele.rowe@education.gsi.gov.uk

This document is also available from our website at:

<http://www.education.gov.uk/researchandstatistics/research>