Department for Education

Study of Early Education and Development (SEED): Impact Report on Early Education Use and Child Outcomes at Key Stage 2

Research report
July 2025

Authors: Joseph Crowley, Charles Wilson, Lovisa Moller Vallgarda, Svetlana Speight and Sehaj Bhatti, National Centre for Social Research

Government Social Research

Contents

List of tables

Acknowledgements

The authors would like to thank the children, parents and teachers who have contributed to the collection of the data to the Study of Early Education and Development (SEED). Without your regular contributions by responding to surveys over the years, this research would not be possible. We are very grateful for your support.

We are grateful to colleagues in different departments at NatCen who were involved in data collection and data management, in particular, Kevin Palmer, Pamela Ritchie, Jan Read, Sonia Shirvington, Moira Hamill, Anna Keyes, Migle Aleksejunaite, Jordan Fenlon, Chung Ho Fung and 170 research interviewers.

The authors would also like to thank staff in the Early Years Analysis and Research Unit at the Department for Education, the SEED Advisory Board as well as Sergei Plekhanov and Marian Morris at SQW, for their advice throughout the research process.

Finally, the authors would like to thank Edward Melhuish and Julian Gardiner for the research design and variable derivation that they produced for the analysis of Key Stage 1 results for this longitudinal study, which this report draws heavily upon, as well as for their helpful clarifications of their research decisions made.

The advisors and reviewers listed above provided constructive comments and suggestions. However, they were not asked to endorse the content or conclusions of this report. The authors alone are responsible for the final content.

This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners.

Executive summary

Key points

- Increased hours of ECEC in formal group settings (nursery classes, nursery schools, or playgroups) was associated with a slight improvement in meeting expected standards in reading, writing, and maths (combined) and reading at KS2.
- An additional 10 hours per week in formal group childcare during early years increased the likelihood of achieving the expected standards in reading, writing, and maths (combined) by approximately 3 percentage points.
- The likelihood of achieving the higher standard in reading, writing and maths (combined) also increased slightly with increased hours spent in formal group settings: the odds increase by 2% with each additional hour a week the child attended.
- Increased hours of ECEC in a formal domestic setting (such as from childminders) was positively associated with achieving the expected standard in science at KS2, whilst increased hours in informal settings (such as from relatives, friends, neighbours) was associated with improved attainment in maths.
- Children from financially disadvantaged families show a more pronounced benefit from increased formal group ECEC attendance. For children who joined the SEED study in the most financially disadvantaged 20% of families, each additional hour of formal childcare per week during their early years was associated with a 4.5% increase in the odds of achieving the expected standard in reading, writing and maths (combined). This improvement is around twice as large as for children overall.
- The positive effects of childcare for financially disadvantaged children was also established at KS1 (Melhuish and Gardner, 2021) and have persisted throughout primary school. Combined, these findings indicate that early years childcare is likely of particular importance for children from disadvantaged backgrounds in helping them to succeed academically.
- A better Home Learning Environment (HLE) index score was linked to a slightly higher likelihood of meeting expected standards in reading, writing, and maths (combined) at KS2.
- An authoritative parenting style, characterised by high responsiveness and high levels of psychological control, and clearer parental limit setting was also found to have a positive impact on KS2 attainment.
- Home environment factors, both in early childhood and later in primary school, were also strongly associated with socio-emotional outcomes at KS2 and able to

explain a greater share of children's socio-emotional outcomes than early years childcare experiences.

- In particular, parental psychological distress, a chaotic home life, a more invasive relationship between mother and child (reflecting issues such as the mother feeling in conflict with or annoyed by her child) were all associated with poorer socio-emotional outcomes at KS2. Associations between these elements of the home environment and poorer outcomes was also found at KS1 (Melhuish & Gardiner, 2020).

Introduction

The analysis presented in this report is based on the Study of Early Education and Development (SEED) survey data. SEED is a major longitudinal study aimed at evaluating the long-term impacts of ECEC and the potential benefits of government investment in this area. Tracking nearly 6,000 children from age two since 2013-2014, the study has been extended to 2029 which will monitor their progress until they reach Year 11 in secondary school.

This report explores whether factors captured in these children's early years by the SEED surveys are associated with their attainment and socio-emotional outcomes upon their completion of KS2 at Year 6. The main focus of this report is how these outcomes are associated with children's experience of early childhood education and care (ECEC), including the amount of time spent in care, the age they started receiving care, and the quality of childcare. The influence of the child's early home environment, such as their parental relationships, is also looked at as a secondary factor.

Over the past few decades, there has been a growing global focus on ECEC to enhance social and educational outcomes for children and reduce inequalities. As a result, both demand for and access to childcare have increased significantly in developed countries. Previous research shows that ECEC can positively influence cognitive, language, behavioural, and social outcomes, particularly for disadvantaged children if the care is of high quality (Berger et al., 2021; Herbaut et al., 2024; Melhuish et al., 2015; Melhuish and Gardiner, 2023; Sylva et al., 2012).

A previous SEED report conducted when the children were around age seven at the end of Key Stage 1 (KS1) explored similar issues. That report, authored by Melhuish and Gardiner in 2021, found that attending higher quality formal group settings (such as nursery classes, day nurseries and pre-schools) was associated with better academic results for maths and science. However, some socio-emotional issues, such as externalising behaviour,[1] were associated with increased use of formal group ECEC (Melhuish & Gardiner, 2020). In this previous SEED report, the home environment and

[1] Externalising behavioural problems refers to outward behaviours, such as being disruptive, hyperactive, or aggressive.

parenting practices were also found to be consistently influential on children's socio-emotional development and attainment, including through a positive association between limit setting from parents and better outcomes overall at KS1.

The analysis presented in this report sought to address the following research questions:

1. How is early childhood education and care linked to attainment at KS2?
2. How is early childhood education and care linked to socio-emotional outcomes at KS2?
3. How is home environment and the quality of the parent/child relationship linked to attainment at KS2?
4. How is home environment and the quality of the parent/child relationship linked to socio-emotional outcomes at KS2?

Data and methods

The analysis presented in this report is based on survey responses from the SEED study and data from the National Pupil Database (NPD). SEED survey data from survey waves 1, 2, 3 and 7 were matched with NPD record extracts to enable the analysis[2].

When the association between ECEC and these outcomes was analysed, three different aspects of their ECEC exposure were considered. The primary focus was on the average number of hours of childcare the children attended up until they started school. The analysis considered different types of childcare separately, grouped into formal group childcare (such as nursery classes, day nurseries and pre-schools), formal individual childcare (childminders) and informal individual care (such as from relatives, friends, neighbours and nannies). In addition to this primary focus, the quality of formal group childcare and the age at which the child first attended group formal childcare (if at all) were considered.

To explore the association with the home environment in early years a range of measures were used. These are all based on survey questions to the parents and designed to capture information about the home environment and the quality of the parent/child relationship. The nine measures considered were the Confusion, Hubbub, and Order Scale (CHAOS), how far parents put limits on their children's behaviour, parental psychological distress, the two Mother Object Relations Scales (MORS) subscales (invasiveness and warmth), the authoritarian parenting score, the authoritative parenting score, the permissive parenting score and the Home Learning Environment (HLE) index.

[2] Waves 4, 5 and 6 were not included in the analysis for this report because the available sample size would be substantially restricted by including data from these waves, in particular with waves collected during the Coronavirus pandemic.

To assess children's academic performance at KS2, the analysis focused on whether the children met the expected standard in reading, writing, and maths (combined); whether they achieved the expected standard in all these subjects and whether the children achieved the higher expected standard and their attainment in specific subjects was also assessed. Children's socio-emotional outcomes at KS2 were measured using the parents' responses to the Strengths and Difficulties Questionnaire (SDQ), using both the SDQ Total difficulties score and the SDQ subscale scores.

A broad range of socio-demographic controls were included in the main models. All analysis was carried out using survey weights.

Early childhood education and care (ECEC) and attainment

Almost all children in the SEED sample attended at least some form of ECEC. Because of this the analysis considers how far an increase in the average hours a child spends in childcare is associated with being more likely to succeed academically at KS2. In that sense they are the effects associated with variations in how much and the type of childcare; not the result of attending or not attending childcare.

Time spent in formal group childcare

The findings show that children who spent more time in ECEC formal group settings like nursery classes, nursery schools, or playgroups between the ages of 2 and 4 were slightly more likely to meet the expected standards in KS2 assessments in reading, writing and maths (combined score). This was the case even when taking other childcare arrangements and a broad range of home environment and socio-demographic factors into account, such as the parents' highest academic qualification and the child's special educational needs (SEN) status.

However, the effect associated with additional time in these childcare settings is modest. The average likelihood of achieving the expected standard in reading, writing and maths at KS2 for children in our sample is 61%. If this probability is adjusted to reflect the effect of an extra 10 hours per week in formal group childcare this increases the chance of meeting the expected standard in by about 3 percentage points, although there is a more noticeable effect for children from the most financially disadvantaged families. When children in the most disadvantaged 20% of households are analysed separately, the positive association between increased childcare hours and attainment is stronger compared to children from other income groups. This is consistent with findings from previous analysis conducted at KS1 (Melhuish and Gardner, 2021). These findings indicate that early years childcare is likely to be of particular importance for children from disadvantaged backgrounds in helping them to succeed academically.

In addition, the association was not consistent across individual subject areas. The results above relate to how likely children are to achieve the expected standard *in all*

three subjects; reading, writing and maths. When the individual subjects were analysed, increased hours of formal group childcare was only found to be positively associated with achieving the expected standard in reading, not with the other subjects.

Time spent in individual childcare

For domestic settings, there was no significant link found between meeting the expected standards in KS2 assessments in reading, writing and maths (combined) and time spent in childcare once other factors were considered. That was the case for both informal individual settings, such as care provided by relatives, friends or neighbours, and formal individual settings, such as care provided by childminders.

Early childhood education and care (ECEC) and socio-emotional outcomes

No connection was found between socio-emotional outcomes at KS2 and the different measures of ECEC which were explored in this report, when controlling for a child's socio-demographic background and their home environment experience during their early years. In earlier SEED reports, an association between children's socio-emotional wellbeing at KS1 and ECEC had been found (Melhuish & Gardiner, 2020). That this association was no longer statistically significant at KS2, using similar controls, suggests that this association was short-term. Further modelling which also considered early years socio-emotional wellbeing produced inconclusive results.

The home environment

The home environment is a key factor in children's academic and socio-emotional development, as demonstrated by its influence on children's outcomes in SEED, both in their early years (Melhuish and Gardiner, 2020; Melhuish and Gardiner 2018), at Key Stage 1 (KS2) (Melhuish and Gardiner, 2021), as well on Key Stage 2 outcomes as described in this report. In the analysis of children's KS2 outcomes in SEED for this report several measures of the home environment were found to be associated with how children were doing at KS2, for both academic and socio-emotional outcomes. In particular, the home learning environment and how far parents set limits on their children's behaviour were found to be important across both of these areas. Others, including how 'authoritarian' a parenting style parents used, the level of disorder in the household and the quality of the mother-child relationship, were found to be associated with only one area. That these factors, which were measured during children's early years, continued to explain children's outcomes at KS2 suggests they have a long lasting-role in children's development.

Home environment at attainment

Focussing on attainment, the analysis found that the nine home environment measures in the study collectively explained a small amount of the variation in children's attainment outcomes at KS2. Of the nine measures of home environment explored, three were found to be associated with achieving the expected standard in reading, writing and maths (combined) at KS2:

- The analysis found the Home Learning Environment (HLE) index to be of consistent importance. Of all the home environment measures, it was the measure most often associated with the different outcomes of academic achievement. These findings show that children with a better home learning environment were found to be more likely to achieve the expected standard in reading, writing and maths at KS2.
- Additionally, a higher score on the authoritative parenting style (more responsive and higher levels of psychological control) was linked to a increased likelihood of meeting the expected standard in this group of subjects.
- The analysis also found that stricter parental limits on children's behaviour were somewhat associated with better outcomes in this group of subjects at KS2.

Home environment and socio-emotional outcomes

The home environment during a child's early years was linked to their socio-emotional outcomes at KS2. Collectively, the nine early years home environment factors captured in the SEED study were able to explain 14.3% of the differences in overall emotional difficulties children faced at KS2. The same variables were able to collectively explain 9.7% of the differences in positive social behaviour, estimated using the prosocial score at this age. When home environment factors from both early years and KS2 together were used to explain socio-emotional outcomes at KS2, these could account for 25.6% of the variation in children's total difficulties and 14.0% of the variation in their positive social behaviour.

Of the nine measures of a child's early years home environment which were explored in this report, four were found to be associated with socio-emotional outcomes at KS2:

- If parents experienced higher levels of psychological distress, such as symptoms of depression or anxiety, in the child's early years children were found to be more likely to have a problematic difficulty score at KS2.
- Parents that reported issues such as struggling to stay on top of things or lacking calm in their home, as measured by the CHAOS scale, were more likely to have children that experienced emotional difficulties.

- Children whose mothers had a high score on the MORS invasiveness subscale, which measures how much the mother felt her child is demanding or irritating, were also more likely to experience socio-emotional difficulties at age 10 to 11.
- Finally, higher levels of parental limit setting during the early years was also associated with a higher likelihood of problematic difficulties scores at KS2.

Introduction

Background and Policy Context

In recent decades, there has been a rise in global interest in early childhood education and care (ECEC), reflecting an increase in the desire to improve social and educational outcomes for children and to reduce inequalities. As such, both demand for and access to childcare have dramatically expanded across developed countries. For most aspects of development, the evidence indicates that ECEC is associated with beneficial effects for children. Research findings show that ECEC can have a positive effect on the cognitive, language, behavioural and social outcomes of children, in both the short-term and the longer-term, particularly for children from more disadvantaged backgrounds and if it is of good quality (Berger et al., 2021; Herbaut et al., 2024; Melhuish et al., 2015; Melhuish and Gardiner, 2023; Sylva et al., 2012).

Within England, following the evidence from the Effective Pre-school, Primary and Secondary Education (EPPSE) study of the positive effects of ECEC upon children's development (Sylva et al., 2004) as well as international evidence, successive governments introduced and expanded provision of funded ECEC to pre-school children of different ages (Melhuish, 2016; HM Treasury, 2023), increasing the number of hours available, reducing the age when children become eligible for the provision and broadening eligibility. Currently, there are the following entitlements available to parents of pre-school children:

- the 15 hours entitlement for eligible working parents of children from 9 months to 2 years old (this is a new entitlement from 1 September 2024; it is 15 hours per week over 38 weeks a year but the hours could be used flexibly depending on arrangements at the provider)
- the 15 hours entitlement for eligible working parents of 2-year-old children (new entitlement from 1 April 2024)
- the 15 hours entitlement for disadvantaged 2-year-olds
- the universal 15 hours entitlement for all 3- and 4-year-olds
- the additional 15 hours entitlement for eligible working parents of 3- and 4-year-olds.

While studies with different populations documented various benefits to children from attending high quality ECEC, it has also been noted that effects of a child's socio-economic background and the home environment and parenting they experienced in the early years are typically more wide-ranging and bigger in size. Household income, maternal education, parental mental health, the quality of the parent-child relationship and the home learning environment (home learning activities with the child) have been found to be associated with a range of cognitive and behavioural outcomes (Gutman et al., 2019; Kiernan et al. 2024; Sammons et al., 2008; Tamura et al., 2020). Some evidence suggests that these factors do not function alone but interact with each other.

Hence the potential effects of ECEC experience may be partly moderated by family factors, such as disadvantage and the home learning environment (Melhuish and Gardiner, 2023).

Study of Early Education and Development (SEED)

The Study of Early Education and Development (SEED) is a major longitudinal study designed to provide evidence on the effectiveness of ECEC and to identify any short- and longer-term benefits from government investment in this provision. The SEED study has been following nearly 6,000 children from when they were aged 2 years (in 2013-2014). The study has been extended to continue following progress of the young people until they are in Year 11 in secondary school (aged 15 to 16 years). The extension study is being conducted by a consortium including the National Centre for Social Research (NatCen), University College London, Durham University, University of Bristol and SQW.

The original SEED study was funded by the Department for Education to explore the impact of providing funded hours of ECEC to disadvantaged two-year-olds on take-up of the provision by eligible families and to assess the associations between ECEC experience and home environment and child development. Since the first impact report was published in 2017 (Melhuish et al., 2017a), the SEED study has continued exploring associations between ECEC and home environment and a range of outcomes related to children's cognitive development, socio-emotional development, and how well they progress at school.

The last SEED report (2021) looked at outcomes at the end of Year 1 (Phonics test pass/fail) and Year 2 (achieving expected level in reading, writing, maths and science), when children were aged 5 to 6 years and 6 to 7 years respectively. This outcome data was matched from records in the National Pupil Database, and there was no survey data collection at that time. Findings from that report show that overall, there were no significant associations between children's Phonics and Key Stage 1 outcomes during school years 1 to 2 and the amount of ECEC received by children between age 2 years and the start of school. However, attending higher quality formal group settings (such as nursery classes, day nurseries and pre-schools) was associated with better academic results for maths and science. Additionally, for disadvantaged children, an early start with a higher amount of formal ECEC between age 2 years and the start of school was associated with better child outcomes (Melhuish and Gardiner, 2021).

Previous SEED reports explored the relationship between children's experiences of ECEC and their cognitive, educational and socio-emotional outcomes at ages 3, 4 and 5 years. This used outcome data collected in SEED surveys at ages 3 and 4 years, and a mix of survey and NPD data at age 5 years. They found a consistent positive association between being looked after by a family member (informal individual ECEC) and verbal ability (Melhuish et al., 2017a; Melhuish and Gardiner, 2018; Melhuish and Gardiner, 2020) and less consistent positive associations between experiences of formal ECEC

and cognitive development (Melhuish et al., 2017a; Melhuish and Gardiner, 2018). Therefore, the findings in relation to socio-emotional development were mixed, with some effects of ECEC on outcomes being positive and others negative.

Reflecting on differences in the findings between the earlier EPPSE study and the SEED reports, Melhuish and Gardiner (2021) suggested that children's ECEC experiences had levelled up in the last two decades, and there was a near universal use of ECEC when the SEED study collected its data on the take-up. There had also been an increase in overall ECEC quality (Melhuish and Gardiner, 2017b), so children's ECEC experiences across the population were more equivalent than before. A potential consequence of this levelling up of ECEC experiences is that any effects of ECEC differences upon child development were likely to be reduced (Melhuish and Gardiner, 2021).

In addition to focusing on the effects of ECEC, SEED reports have also explored the significance of the home environment. The report at age 7 found that children who grew up in families with a richer home learning environment (where parents engaged in home learning activities with the child more frequently), who had a warmer relationship with their parents and whose parents were more likely to set limits achieved better outcomes overall at Key Stage 1. Conversely, permissive parenting was associated with poorer child performance on these outcomes. For the Phonics check, there were significant associations with home learning environment and warmth in the parent-child relationship (Melhuish and Gardiner, 2021).

In the SEED reports looking at cognitive and educational outcomes at ages 3, 4 and 5 years, the home learning environment and limit setting proved to have the strongest associations with children's cognitive development and progress at school. Other home environment factors that were significantly associated with children's better cognitive and educational outcomes (although not as consistently as home learning environment) were closeness between the parent and the child and warmth of the relationship. Permissive or authoritarian parenting, parental psychological distress and chaotic environment at home were found to be negatively associated with cognitive and educational outcomes (Melhuish et al., 2017a; Melhuish and Gardiner, 2018; Melhuish and Gardiner, 2020).

As regards socio-emotional outcomes, which were measured at ages 3, 4 and 5 years, children who grew up in chaotic households and whose parents experienced psychological distress tended to have poorer outcomes, and those who had closer and warmer relationships with their parents tended to have better outcomes. Parental limit setting was also significant for socio-emotional outcomes, but the results were mixed, not always indicating benefits for the child (Melhuish et al., 2017a; Melhuish and Gardiner, 2018; Melhuish and Gardiner, 2020).

Objectives of this report

The objectives of this report were to answer the following research questions:

1. How is early childhood education and care linked to attainment at Key Stage 2?
2. How is early childhood education and care linked to socio-emotional outcomes at Key Stage 2?
3. How is home environment and the quality of the parent/child relationship linked to attainment at Key Stage 2?
4. How is home environment and the quality of the parent/child relationship linked to socio-emotional outcomes at Key Stage 2?

Data and methods

The Study of Early Education and Development (SEED)

The analysis presented in this report is based on survey responses from the Study of Early Education and Development (SEED) survey waves 1, 2, 3 and 7. This a longitudinal study that has followed nearly 6,000 children from the age of 2. In this analysis most of the predictors come from Waves 1 – 3, which were conducted in 2013 – 2016 (ages 2 to 4). Whereas the outcomes are measured at wave 7, conducted in 2022 and 2023 (ages 10 to 11). In wave 3 the responding sample size was 3,930, in Wave 7 it was 2,146. Data from the intervening years (conducting during the Coronavirus pandemic) are excluded to preserve sample sizes as the response rate in these waves was lower.

The SEED study sample was drawn to be representative of families in the population with children aged two in England, at the time of the original survey (2013). It included an oversampling of families from more disadvantaged backgrounds to ensure there were large enough samples in these groups for them to be analysed separately. All analysis was conducted using survey weights that adjust for non-response and unequal sampling probabilities, to ensure the results remained representative of the original population from which it was drawn.

Outcomes

Two outcomes of interest are analysed in this report: academic attainment and socio-emotional wellbeing. The main estimate of children's socio-emotional outcomes at KS2 comes from the Strengths and Difficulties Questionnaire (SDQ) that parents completed about their child at KS2 (SEED wave 7). The total difficulties score was selected as the main outcome, providing an overall measure of wellbeing, with the subscales of the SDQ treated as secondary outcomes. The subscales measures problems with conduct, emotions, behaviour and peers, as well as the presence of prosocial behaviours (see Appendix 2 for full description). The sample size for this analysis, which relies on data from wave 7, is approximately 2,140 – the total responding sample size at Wave 7.

To explore children's academic attainment at Key Stage 2 (KS2) data from the National Pupil Database was used. This includes a wide range of administrative data about children's progress through the education system. In this analysis data about children's performance at KS2 was used, with a binary indicator of whether or not a child had achieved the expected standard in reading, writing, and maths at KS2 adopted as the main attainment outcome of interest. This a single measure of attainment, indicating whether someone achieved the expected standard in all these subjects. A number of secondary measures of attainment were also analysed (see Appendix 2), in the body of the report the commentary is restricted to achievement of the higher-than-expected

standard at KS2 – in all three of the subjects reading, writing and maths. The sample size in this analysis was greater, (N=3,930) because only predictors from Wave 3 were used and no outcomes collected at Wave 7 were required (which would have limited the sample size to the approximate two thousand cases responding in that wave).

Predictors

Early childhood education and care

The analysis included two groups of predictors. The first set are measures of early childhood education and care (ECEC) use. The primary variables for this are:

- Hours of informal childcare at age 0 to 4 (weekly average across waves).
- Hours of formal group childcare at age 0 to 4 (weekly average across waves).
- Hours of formal individual childcare at age 0 to 4 (weekly average across waves).

Two additional measures of ECEC use are also included.

The first is the age a child first received 10 hours or more a week (on average) of formal early childhood education and care (ECEC). This includes both hours of formal group and formal individual childcare (treated as a categorical variable).

The second, collected for 1,465 children within the sample, is an assessment of the quality of the group, formal childcare provision a child received (treated as a numeric score), called the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale. See Appendix 2 for more details of these.

It should be noted that less than one per cent of the children in the SEED study did not attend any ECEC provision before they started school (see Melhuish & Gardiner 2021, 'Appendix 2: ECEC setting quality measure', for further details). The research design is therefore focused on establishing the effects associated with an increase or decrease in ECEC attendance. The research covers the effects associated with different types of ECEC provision, the number of weekly hours attended, and the quality of provision. It does not explore the effects of attending vs. not attending ECEC, as a control group is not present.

The home environment

The second set of predictors are early years home environment factors. They focus on the nature of the parenting the child received and the quality of the learning environment within the home during the child's early years. These include the following (see Appendix 2 for full details):

- Confusion, Hubbub, and Order Scale (CHAOS) (average for W1 & W2).
- How far parents put limits on their children's behaviour (average for W1 & W2).
- Parental psychological distress (average for W1 & W2).
- Mother Object Relations Scales (MORS) invasiveness score (W2).
- Mother Object Relations Scales (MORS) warmth score (W2).
- Authoritarian parenting score (W3).
- Authoritative parenting score (W3).
- Permissive parenting score (W3).
- Home Learning Environment (HLE) index (average for W1, W2 and W3).

Modelling approach

The measures of academic attainment were binary (that is to say, they took only yes / no values). For these outcomes binary logistic regression models were used in this analysis. The SDQ outcomes for socio-emotional wellbeing are numeric scores. For these linear fixed effects models were applied (see Appendix 1 for further details).

Covariates

A range of socio-demographic control variables are also included in the models presented in this report. These are ethnicity, sex, SEED sample deprivation group, experience of adverse events, whether the child had a health, development, or behavioural issue, the parents' highest qualification, the SDQ total difficulties score (Wave (W) 2-3), SEN status, FSM receipt and CIN status (see Appendix 2 for further details).

Chapter 1: Early childhood education and care (ECEC) and attainment at Key Stage 2

Key points

- Increased time spent in early childhood education and care (ECEC) in nursery classes, nursery schools or playgroups (formal group childcare) between ages 2 and 4 was associated with a slight increase in the likelihood of achieving the expected standard in reading, writing, and maths (combined) at the end of Key Stage 2 (KS2).
- The effect associated with an increase in formal group childcare hours is modest, although statistically significant. The effect size of a 10-hour increase in formal group childcare equates to around a 3 percentage point increase in the probability of children achieving the expected standard. The effect size is larger for children in the most financially disadvantaged families.
- The association was not consistent across individual subject areas. Increased hours of formal group childcare were found to be positively associated with achieving the expected standard in reading at KS2, when controlling for socio-demographic factors and hours spent in other forms of childcare. No significant association was found in other subjects.
- No association between likelihood of achieving the expected standard in reading, writing, and maths (combined) and attending ECEC in a domestic setting (e.g. from relatives, friends, neighbours and nannies and childminders) was found, when socio-demographic factors and other forms of childcare attended were also controlled for.
- For individual subjects, hours of informal childcare were found to be positively associated with achieving the expected standard in maths and hours of formal individual childcare were found to be positively associated with achieving the expected standard in science.

Attainment and childcare hours

The analysis explored the association between the amount of ECEC children received between ages 2 and 4 years and their attainment at age 10 to 11 years, as captured by the KS2 assessments. The main outcome of interest is whether a child reaches the expected standard in all three of the subjects reading, writing and maths. A child's likelihood of reaching the higher expected standard in these three subjects is also explored.

Table 1: Regression results for overall attainment and childcare hours

	(1) Expected standard	(2) Higher standard
Hours of informal childcare	1.004	1.005
Hours of formal group childcare	1.015*	1.020*
Hours of formal individual childcare	1.003	1.003

Note: The results in the table are odds ratios. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 42.

The results presented in model 1, Table 1 indicate a positive association between increased hours of formal group childcare (in nursery classes, nursery schools or playgroups between ages 2 and 4) and a slight increase in the likelihood of achieving the expected standard in the reading, writing and maths assessments at the end of KS2. The odds of achieving the expected standard associated with each additional childcare hour per week is 1.5% (odds ratio 1.015; model 1), controlling for socio-demographic factors, the home environment, and hours in other forms of childcare attended (if any). There was no significant association between attainment and increased childcare attendance in a formal individual (childminder) or informal setting (relatives, friends, neighbours or nannies).

Model 2 in Table 2 explores the relationship between childcare hours and achieving the higher expected standard. A positive association between increased hours of formal group childcare and attainment at KS2 is found also for this outcome. The odds of achieving the higher standard increased by 2% with each additional childcare hour per week that the child attended (odds ratio 1.020; model 2), controlling for socio-demographic factors, the home environment, and hours in other forms of childcare attended (if any). No similar association between formal individual childcare or informal childcare is established.

The full model results are included in the appendix (Appendix 6: Complete results chapter regression tables: Table 42), along with a model where socio-demographic predictors alone are used to explain attainment outcomes ('Appendix 3: Selection of socio-demographic control variables'). These results show that childcare hours, home environment factors and sociodemographic factors combined are marginally better at explaining attainment outcomes than a model with just socio-demographic predictors.

An interaction term between special educational needs (SEN) status and formal group childcare hours was also explored. The relationship between additional formal group

childcare hours (in nursery classes, nursery schools, or playgroups) and the likelihood of achieving the expected standard in reading, writing and maths, was not found to differ between children with SEN and children without SEN. In other words, increased hours of ECEC in formal settings was associated with a slight improvement in meeting expected standards in reading, writing, and maths (combined) and reading at KS2 for both children with and without SEN.

Attainment and childcare hours, by financial disadvantage

The SEED longitudinal survey benefits from having a large sample of respondents whose families qualified for funded early education entitlement for two-year-old children, as their families were in receipt of specified benefits. These families, who might be regarded as the most financially disadvantaged, are oversampled by design in the SEED study (see Appendix 1). The oversampling of children from the most disadvantaged families enables analysis by level of disadvantage.

For children of the most financially deprived third of families in the study, having attended additional hours of group formal childcare appears to have a greater impact. For these families, there is a 4.5% increase in the odds associated with each additional weekly childcare hour in a formal group setting (odds ratio of 1.045), when controlling for socio-demographic factors and other childcare hours (see Table 2, Model 2). No similar association is identified for the intermediate group or the least deprived group, when considered separately. These findings indicate that, whilst among the most financially disadvantaged people the likelihood of achieving the expected standard increases with the amount of time they spend in ECEC, this is not the case for other children.

However, it is worth noting that these subgroup analyses have smaller sample sizes than the previous models. Because they have smaller samples it is less likely that results will be statistically significant.

Table 2: Regression results for overall attainment and childcare hours, stratified by disadvantage

	(1) Full sample	(2) Most deprived	(3) 2nd quintile	(4) Least deprived
Hours of informal childcare	1.007	1.021	1.007	1
Hours of formal group childcare	1.017*	1.045*	1.001	1.020+
Hours of formal individual childcare	1.004	1.046	0.995	1.005

Note: The results in the table are odds ratios (a value below 1 indicates a decrease in the odds of experiencing that outcome; a result above 1 an increase). Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and ***

denotes p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 43.

Attainment in individual subjects and childcare hours

To gain a more complete picture of the relationship and to enable further modelling, models of the likelihood of a child achieving the expected standard or higher expected standard in individual subjects (those covered by the national curriculum tests and the teacher assessment of each child's English writing and science) were produced. These are available in 'Appendix 4: Further attainment models'.

In these further models, hours of formal group childcare were found to be positively associated with achieving the expected standard in reading at KS2. When controlling for socio-demographic factors and hours spent in other forms of childcare each additional hour of formal group childcare was associated with a 1.9% increase in the odds of a child achieving the expected standard in reading at KS2 (Table 32).

Hours of informal childcare were found to be positively associated with achieving the expected standard in maths at KS2. When controlling for socio-demographic factors and hours spent in other forms of childcare, each additional hour of informal childcare was associated with a 1.2% increase in the odds of a child achieving the expected standard in maths at KS2 (Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in grammar, punctuation, and spelling at KS2 (model 1) and whether the child had achieved the higher standard in grammar, punctuation, and spelling at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 31).

Hours of formal individual childcare were found to be positively associated with achieving the expected standard in science at KS2. When controlling for socio-demographic factors and hours spent in other forms of childcare, each additional hour of formal individual childcare was associated with a 3.9% increase in the odds of a child achieving the expected standard in science at KS2 (Table 34).

Attainment and age started formal ECEC

For all children, the age at which they started formal group childcare (if at all) is calculated based on the age they first attended 10 hours a week (on average) of formal childcare.

As would be expected, the age when a child starts formal group childcare and their average weekly use over the years are highly related concepts. In other words, children in the study who attended more hours of group formal childcare before they started primary school also tended to start childcare earlier. Table 4 displays a crosstabulation of the two childcare use variables. It shows that an estimated 76% of children who start nursery before age 1 will also be in the high use group (the highest quartile of childcare hours per week on average). Only an estimated 1% of children who start nursery after age 4 are in the high-use group.

The analysis considered whether differences in when a child attended at least 10 hours per week formal ECEC were associated with attainment at the end of primary school. Table 5 presents the results. Model 1 in the table displays modelled odds ratios for achieving the expected standard and model 2 displays modelled odds ratios for achieving the higher expected standard in reading, writing and maths at KS2, both controlling for socio-demographic and home environment characteristics.

The odds ratios show the change in the odds of successfully attaining these academic results, compared to people who never received more than 10 hours a week of ECEC. Most of the odds ratios in these models are both quite high and positive, which indicates the people who started receiving ECEC for more than 10 hours a week before age 4 are more likely to succeed academically. However, the large standard errors for the estimates mean that these results are not statistically significant. That is to say, we cannot be confident they are not simply due to random sample variation.

The only finding which approached statistical significance was in model 1: starting nursery between age 1 and 2 was found to be associated with an 106% improvement in the odds (odds ratio 2.061) of achieving the expected standard, compared to children who never received more than 10 hours a week of formal ECEC. However, to reiterate this result was only statistically significant at the 10% level.

Overall, the models of average hours of childcare attended per week discussed above appears to give us a somewhat better understanding of the relationship between childcare use and attainment than when age started is used as the indicator.

Attainment and quality of provision

A subset of children also had the quality of their formal childcare provision assessed. In this analysis this is measured by the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale, a measure of the quality of staff interactions with children in the setting. Details of this can be seen in Appendix 2: Variable descriptions.

To assess the quality of the formal childcare some children were receiving, 1,000 formal childcare providers were assessed that children within the SEED study attended. This was not a random sample of childcare providers but was selected to keep the proportion of settings of different types (nursery class, children's centre etc) roughly equivalent to the proportion attending these types of setting in the wider SEED sample. In the Wave 3 sample used in this analysis 38% of the full sample had a quality measure (1,465 children).

Table 6 displays the results. There is no statistically significant association found between the ECEC quality assessment and a child's attainment in reading, writing and maths (combined) at KS2, either for achieving the expected standard (model 1) or the higher standard (model 2).

This is in line with the descriptive statistics of the association between achieving the expected standard in reading, writing, and maths at KS2 and the childcare quality measure (SSTEW) displayed in Table 3. This crosstabulation, without controlling for other variables, displays that 60% of children who attended a setting classed as either minimal or inadequate in the early years achieved the expected standard, compared to 62% of children who attended a setting classified as good and 60% of children who attended a setting classed as excellent (no statistically significant differences).

Previous research had identified a positive association between quality of early years childcare and improved outcomes in several areas of cognitive development in their early years (Sammons et al., 2002), although in earlier waves of the SEED study this was not found to be the case (Melhuish & Gardiner, 2020). Potential factors which may have a role in explaining why no statistically significant association between setting quality and attainment was found in these results are that a) the sample size for the quality analysis is smaller, b) more children now attend ECEC such that in the SEED study almost all children have attended some formal group ECEC, and c) the general quality of ECEC has improved markedly over recent years (Melhuish & Gardiner, 2020). In the Effective Pre-school, Primary and Secondary Education (EPPSE) study, conducted in 1997-2000, the average score on a measure of overall ECEC setting quality increased from 4.29 to

5.18, a shift from an ‘adequate’ to a ‘good’ rating (on the Early Childhood Environment Rating scale) (Melhuish and Gardiner, 2017b).

Discussion and conclusion

The slight association between increased childcare attendance and academic achievement in school, observed in the Key Stage 1 (KS1) report (Melhuish and Gardiner, 2021), is also evident in this KS2 analysis. The KS1 report found that children who attended higher-quality ECEC in nursery classes, nursery schools, or playgroups before starting school perform slightly better in KS1 Maths, KS1 Science, and in a combined KS1 English and Maths assessment during their second year of school (ibid.).

At Key Stage 2 (KS2), this study found that children who attended ECEC in formal settings such as nursery classes, nursery schools, or playgroups were marginally more likely to meet the expected standards in reading, writing, and maths (combined). The positive association between increased childcare hours and achieving the expected standard in all three of these subjects held also when taking other childcare arrangements and a broad range of socio-demographic factors, such as the parents’ highest academic qualification, into account.

The impact of additional time in these settings at KS2 is modest. For example, consider a child who otherwise would have the average probability of 61% for achieving the expected standard in our sample for reading, writing and maths. An extra 10 hours per week in formal group childcare on average throughout the early years increases their likelihood of meeting expected standard in English, writing and maths (combined) by approximately 3 percentage points. Turning to individual subject areas, increased hours in formal group childcare were positively associated with achieving the expected standard in reading at KS2, but no association was found in other subjects.

A more pronounced effect is observed for children from families experiencing higher levels of financial disadvantage, which is consistent with findings from KS1 analysis. These analyses suggest that the positive association between increased childcare attendance and academic attainment is likely to be more significant in these families, with the association persisting throughout primary school.

Table 3: Achieved expected standard in reading, writing, and maths at KS2 by SSTEW (mean of W1 & W2) Categorical

Achieved expected standard in reading, writing, and maths at KS2	Total	No setting quality measure	Minimal or inadequate	Good	Excellent
No	39% [37-42%]	40% [37-43%]	40% [34-47%]	38% [34-42%]	40% [32-48%]
Yes	61% [58-63%]	60% [57-63%]	60% [53-66%]	62% [58-66%]	60% [52-68%]
Unweighted base	*3349*	*2088*	*332*	*757*	*172*
Weighted base	*3307*	*2036*	*335*	*757*	*179*

Base: All respondents at Wave 3.

Table 4: Average number of hours of formal group childcare cross-tabulated with the age the child first received at least 10 hours of formal group childcare

Average number of hours of formal group childcare ages 0 to 4 a week	Total	Never received more than 10 hours of formal CC	Before age 1	Between age 1 and 2	Between 2 and 3	Between 3 and 4	After the age of 4
None	1% [1-2%]	44% [33-56%]	1% [0-2%]	1% [0-2%]	0% [0-1%]	0% [0-0%]	*
Low	23% [21-25%]	53% [41-64%]	10% [8-12%]	8% [6-11%]	3% [2-4%]	15% [13-18%]	83% [79-86%]
Lower medium	26% [24-28%]	3% [1-12%]	6% [4-8%]	6% [4-8%]	8% [6-11%]	57% [54-61%]	11% [8-14%]
Higher medium	25% [23-26%]	*	8% [5-11%]	19% [15-23%]	64% [60-68%]	24% [21-27%]	5% [4-7%]
High	25% [23-28%]	*	76% [71-80%]	67% [61-72%]	25% [21-29%]	3% [2-4%]	1% [0-2%]
Unweighted base	*3,930*	*90*	*597*	*480*	*710*	*1,392*	*661*
Weighted base	*3,930*	*90*	*579*	*484*	*706*	*1,440*	*631*

Base: All respondents at Wave 3

Table 5: Regression results for overall attainment and age started formal ECEC

	(1) Expected standard	(2) Higher standard
Before age 1	1.768	1.613
Between 1 and 2	2.061+	1.689
Between 2 and 3	1.727	0.982
Between 3 and 4	1.492	1.384
After the age of 4	1.465	1.37

Note: The results in the table are odds ratios. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 44.

Table 6: Regression results for overall attainment and ECEC setting quality

	(1) Expected standard	(2) Higher stand-ard
Good	1.063	0.986
Excellent	0.848	0.908
No setting quality measure	0.948	1.562

Note: The results in the table are odds ratios. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 45.

Chapter 2: Early childhood education and care (ECEC) and socio-emotional outcomes at Key Stage 2

Key points

- The amount, type and quality of early childhood education and care (ECEC) attendance was not found to be associated with socio-emotional outcomes at KS2, when controlling for a child's socio-demographic background and their home environment experience during their early years.
- This contrasts with earlier SEED reports, which found an increased likelihood of some negative outcomes for children's socio-emotional wellbeing and associated with greater use of formal ECEC during school Year One (Melhuish & Gardiner, 2020). That the findings of this report, which used similar controls to analysis of earlier SEED waves, did not find the same relationship to be statistically significant suggests that this association was short-term.
- Further modelling which also considered early years socio-emotional wellbeing produced inconclusive results.

Socio-emotional outcomes and childcare hours

The analysis explored the association between the amount of early childhood education and care (ECEC) children received between ages 2 and 4 and their socio-emotional outcomes at age 10 to 11, measured by the Strengths and Difficulties Questionnaire (SDQ).

The main outcome of interest is a child's overall score on the SDQ total difficulties score, which measures the degree of problems children have had across the hyperactivity, peer, emotional and conduct problems scales to give an overall measure of mental health.

The prosocial scale in the SDQ, which measures positive behaviours such as being considerate of other people's feelings or sharing with other children, is also covered. While the total difficulties score is a measure of mental health problems with a higher score indicating poorer mental health, the prosocial scale measures positive social behaviours with a higher score indicating more positive responses.

Model 1 in Table 7 explores the association between hours spent in childcare and the child's total difficulties score at KS2. The results indicate that there are no statistically significant associations (at the 95% confidence level) between ECEC and socio-emotional outcomes for children at KS2.

Similar to the findings for the total difficulties score, there is not a statistically significant association between ECEC and the prosocial behaviours score at KS2 (Model 2),

although it is worth noting here that there is an association significant at the 10% level (which indicates that increased hours of formal group childcare could be associated with a slight decrease in prosocial behaviours at KS2).

Table 7: Regression results for socio-emotional outcomes and childcare hours

	(1) Total difficulties score	(2) Prosocial score
Hours of informal childcare	0.014	-0.007
Hours of formal group childcare	0.011	-0.014+
Hours of formal individual childcare	0.017	-0.007

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. All models in the table control for socio-demographics and home environment factors. The early years total difficulties score, which is included among the controls in the attainment models, is not included among the controls in these models. Levels of statistical significance are denoted as: + p <= 0.1, * p ≤ 0.05, ** p ≤ 0.01, and *** p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 46.

Socio-emotional outcomes and age started formal ECEC

Children in the SEED study started childcare at different ages. The analysis considered when each child attended at least 10 hours per week formal ECEC and whether an earlier start had a notable association with socio-emotional outcomes at the end of primary school.

As outlined in chapter 1, the age when a child starts formal group childcare and their average weekly use over the years are highly related concepts. For example, only an estimated 1% of children who start nursery after age 4 are in the high-use group in terms of average weekly hours.

The analysis explored whether differences in when a child first attended at least 10 hours per week formal ECEC were associated with children's socio-emotional outcomes at KS2. The findings, presented in Table 8, show that there was no statistically significant relationship between these two characteristics, after controlling for the socio-demographic backgrounds of children and their home environments in their early years. This finding was consistent for both the SDQ total difficulties score and the prosocial score.

Table 8: Regression results for socio-emotional outcomes and age started receiving childcare

	(1) Total difficulties score (W7)	(2) Prosocial score (W7)
Before age 1	1.577	-0.193
Between 1 and 2	1.651	-0.123
Between 2 and 3	1.478	0.091
Between 3 and 4	1.676	0.075
After the age of 4	1.467	0.05

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. The early years total difficulties score, which is included among the controls in the attainment models, is not included among the controls in these models. Levels of statistical significance are denoted as: + p <= 0.1, * p ≤ 0.05, ** p ≤ 0.01, and *** p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 48.

Socio-emotional outcomes and quality of provision

Analysis was also conducted to assess whether the quality of the formal childcare children received in their early years was associated with their socio-emotional wellbeing at KS2. Table 9 presents the results. No association between the assessed quality of the ECEC a child received and their socio-emotional outcomes (either the total difficulties score or the prosocial score) was found, while holding constant children's socio-demographics and home environment factors.

This finding from the modelling, once socio-demographic and home environment factors have been controlled for, is broadly in line with the descriptive analysis of these outcomes. Table 10 shows that 18% of children in settings with minimal or inadequate quality were found to have a problematic score on the SDQ total difficulties scale, compared to 14% in settings rated good. However, the confidence intervals for these results cross over substantially, indicating that this difference is unlikely to be statistically significant.

Table 9: Regression results for socio-emotional outcomes at KS2 and the setting quality of the ECEC attended at preschool age

	(1) Total difficulties score	(2) Prosocial score
Good	-0.045	-0.223
Excellent	1.293	0.175
No setting quality measure	0.155	-0.043

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. The early years total difficulties score, which is included among the controls in the attainment models, is not included among the controls in these models. Levels of statistical significance are denoted as: + p <= 0.1, * p ≤ 0.05, ** p ≤ 0.01, and *** p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 49.

Table 10: Whether the respondent experienced socio-emotional difficulties, by the quality of their formal group childcare setting

Whether respondent experienced socio-emotional difficulties	Total	No setting quality measure	Minimal or inadequate	Good	Excellent
No	83% [80-86%]	83% [79-86%]	82% [73-88%]	86% [81-90%]	*
Yes	17% [14-20%]	17% [14-21%]	18% [12-27%]	14% [10-19%]	*
Unweighted base	*2102*	*1325*	*207*	*462*	*108*
Weighted base	*2083*	*1320*	*210*	*445*	*109*

Base: All respondents at Wave 7. Presence of socio-emotional difficulties in this table are measured by the total difficulties score.

Socio-emotional outcomes and SEN

Whether a child receives SEN support, has an Education, Health and Care (EHC) plan is included in all the models describe above as a covariate. This means these findings can

be considered as the effect of hours spent in childcare on socio-emotional wellbeing while controlling for special educational needs. To further explore whether hours spent in ECEC has a differential effect for children with SEN, an interaction effect was introduced to each model, in the same way as introduced for the attainment outcomes (see Chapter 1 - Attainment and childcare hours). However, for the main socio-emotional outcomes (the SDQ Total difficulties score and the Prosocial score) none of these were found to be statistically significant. This suggests the time spent in childcare has a similar association with socio-emotional wellbeing for children with SEN and those without.

Controlling for early years socio-emotional wellbeing

Table 11 below shows the association between hours spent in childcare and the child's total difficulties score and prosocial score at KS2, after including an additional control in the analysis – a child's socio-emotional wellbeing during their early years. This was measured by the SDQ total difficulties score collected in waves 2 and 3, the same time period in which children were attending early years childcare.

The analysis was also repeated for each of the four subscales that make up the total difficulties score of the SDQ. **Error! Reference source not found.** displays regression results where the association between childcare hours and each of these subscales is modelled in turn. These focus on conduct problems, emotional problems, hyper-activity problems, and peer problems.

These models could suggest that by including early years well-being in the analysis, the regression model can separate the effects of early mental health from those of childcare, revealing this slight negative association between increased hours in formal childcare and latter overall socio-emotional wellbeing, and conduct problems specifically. This interpretation is, however, complicated by the fact that time spent in early years childcare has been linked in previous analysis of SEED to mental well-being during the early years, and our model does not consider that indirect pathway. The interpretation of these models is discussed in more detail in the Discussion and conclusion section of this chapter.

Table 11: Regression results for socio-emotional outcomes and childcare hours, controlling for socio-emotional wellbeing in early life

	(1) Total difficulties score	(2) Prosocial score
Hours of informal childcare	0.005	-0.006
Hours of formal group childcare	0.032*	-0.017*
Hours of formal individual childcare	0.046*	-0.015+

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. All models in the table control for socio-demographics and home environment factors, as well as early years SDQ experience. Levels of statistical significance are denoted as: + p <= 0.1, * p ≤ 0.05, ** p ≤ 0.01, and *** p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 46.

Table 12: Regression results for the SDQ subscales and childcare hours

	(1) Conduct problems	(2) Emotional problems	(3) Hyper-activity problems	(4) Peer problems
Hours of informal childcare	0.006	0.001	0.001	-0.001
Hours of formal group childcare	0.013**	0.009	0.014+	-0.001
Hours of formal individual childcare	0.017*	0.017	0.017+	-0.004

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. Levels of statistical significance are denoted as: + p <= 0.1, * p ≤ 0.05, ** p ≤ 0.01, and *** p ≤ 0.001. See complete table in Appendix 6: Complete results chapter regression tables: Table 47.

Discussion and conclusion

In the primary models presented in this chapter no association was identified between early years childcare and socio-emotional wellbeing at KS2, after controlling for socio-demographic background and early years home environment. Previous analysis of SEED (Melhuish & Gardiner, 2020) did find a negative connection between ECEC and socio-emotional scales, with a similar set of control variables, in the first year of school. That this effect no longer appears by KS2 in models with similar controls suggests that the earlier association identified was short-term.

In the additional analysis, controlling for early years socio-emotional wellbeing, a different result was found. These associations should however be interpreted with caution, considering the small effect sizes and the timing of the early years SDQ measurement. This measurement at SEED waves 2 and 3 coincides with the time spent in childcare for parts of the sample. As time spent in formal group ECEC has been linked in previous analysis of SEED to a slight reduction in mental well-being at KS1, this association displayed in the models could also be driven by a small increase in difficulties during early years rather than a long-term impact. Given the ambiguity introduced by the partial overlap in timing of the early years SDQ measurement and the time spent in childcare, further modelling was not carried out.

Chapter 3: Home environment and attainment outcomes at Key Stage 2

Key points

- Children with a better home learning environment, with a higher authoritative parenting score, and those whose parents set more limits on their child's behaviour were more likely to perform well academically at KS2.
- Among children in the top fifth of scores on the Home Learning Environment (HLE) index, which measured activities in the home that promote or provide learning opportunities such as a parent reading with their child, 69% achieved the expected standard in reading, writing and maths (combined), compared to 54% in the lowest fifth of scores.
- Among children with parents who reported low parental limit setting, 57% achieved the expected standard at KS2 in reading, writing and maths (combined). This can be compared to 61% among those whose parents reported high parental limit setting.
- Collectively, the set of nine home environment factors explained a small amount of the variation in children's attainment outcomes at Key Stage 2 (KS2).

Achieving the expected standard

The analysis in this chapter explored the association between children's academic attainment at KS2 and nine measures of the child's home environment. The home environment measures include the HLE index, the two MORS subscales (warmth and invasiveness), the CHAOS scale, as well as measures of how far parents put limits on their children's behaviour, parental psychological distress, authoritarian parenting, authoritative parenting and permissive parenting. 'Appendix 2: Variable descriptions' includes a description of each of these measures.

Collectively, the nine home environment measures explained a small amount of the variation in children's attainment outcomes at KS2. This can be established by comparing the goodness-of-fit of a model with and without including measures of early years childcare as predictors (see complete results in Appendix 6: Complete results chapter regression tables: Table 50). This suggests that the collective impact of the nine measures on children's academic attainment is smaller, relative to their socio demographic background.

However, of the nine home environment measures, three were found to be associated with achieving the expected standard at KS2 across reading, writing and maths (combined), while controlling for socio-demographics and hours spent in childcare. These results indicate that a higher authoritarian parenting score, more parental limit setting on children's behaviour, and a better home learning environment, were associated with an increased likelihood of achieving the expected standard. These findings were robust after controlling for a child's socio-demographics and their use of early years childcare (Table 13).

Table 13. Regression results for achieving the expected standard in Reading, Writing, and Maths on children's early years home environment

	(1) Achieved expected standard
How far parents put limits on their children's behaviour (average W1 and W2)	1.200*
CHAOS scale - average for W1, W2	1.003
Parental mental distress (Waves 1 and 2)	0.986
MORS invasiveness subscale (W2)	1.016
MORS warmth subscale (W2)	0.979
Authoritarian parenting score (W3)	0.88
Authoritative parenting score (W3)	1.309*
Permissive parenting score (W3)	0.984
HLE index (average over W1, W2 and W3)	1.022*

Note: The results in the table are odds ratios. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. The predictors are correlated, so individual odds ratios should be interpretated with caution. See complete results in Appendix 6: Complete results chapter regression tables: Table 50.

Because of the high level of correlation observed between these variables, attributing a unique effect on the outcome to each predictor can be difficult. A high level of correlation among variables in a regression model can make the individual coefficients predicted from the model unreliable. It also inflates the standard errors associated with the coefficients, making it difficult to determine the statistical significance of particular predictors.

Taking these limitations of the data into account, rather than report the coefficients themselves, descriptive findings of the bivariate relationship between a home environment predictor and academic attainment are presented for those which were found to be statistically significant in the regression model.

Bivariate associations between the home environment and academic attainment

The relationship between all home environment factors and achieving the expected standard in reading, writing and maths (combined) is explored using bivariate crosstabulations below. The descriptive findings for those coefficients found to be statistically significant in Table 13 are reported below. Tables for results not shown are provided in the separate tables annexe to this report.

The Home Learning Environment (HLE) index was split into quintiles. Among children with the lowest HLE Index scores 54% had achieved the expected standard at KS2 in reading, writing and maths (combined), compared to 69% among children with the highest 20% of HLE index scores and 65% in the fourth quintile (the second highest HLE scores). For the other groups, the differences were smaller and the confidence intervals overlapped, although the general trend towards better academic attainment was present in each quintile (Table 14).

Table 14: Achieved expected standard in reading, writing, and maths at KS2 by HLE index (quintiles - W1 - W3)

Achieved expected standard in reading, writing, and maths at KS2	Total	Lowest 20%	Second quintile	Third quintile	Fourth quintile	Highest 20%
No	39% [37-42%]	46% [41-51%]	45% [40-50%]	38% [34-42%]	35% [32-39%]	31% [27-35%]
Yes	61% [58-63%]	54% [49-59%]	55% [50-60%]	62% [58-66%]	65% [61-68%]	69% [65-73%]
Unweighted base	*3349*	*665*	*699*	*710*	*646*	*629*
Weighted base	*3307*	*663*	*679*	*701*	*638*	*626*

Base: All respondents at Wave 3.

The authoritative parenting score: due to small sample sizes in some groups the findings for the authoritative parenting score are not reported here (Table 15).

Table 15: Achieved expected standard in reading, writing, and maths at KS2 by Authoritative parenting score (banded W3)

Achieved expected standard in reading, writing, and maths at KS2	Total	Low/Medium	High
No	39% [37-42%]	*	39% [36-41%]
Yes	61% [58-63%]	*	61% [59-64%]
Unweighted base	*3349*	*49*	*3235*
Weighted base	*3307*	*51*	*3181*

Base: All respondents at Wave 3.

The parental limit setting score ranges between 0 and 5. Among children who had a score of less than two, 57% achieved the expected standard at KS2 in reading, writing and maths (combined). This rises to 62% among children with a score of between two and three and 61% among those with a score of more than three, although the confidence intervals for all these percentages overlap (Table 16).

Table 16. Achieved expected standard in reading, writing, and maths at KS2 by How far parents put limits on their children's behaviour (bands W1 and W2)

Achieved expected standard in reading, writing, and maths at KS2	Total	Less than 2	Two to three	Three to five
No	39% [37-42%]	43% [39-48%]	38% [35-41%]	39% [36-42%]
Yes	61% [58-63%]	57% [52-61%]	62% [59-65%]	61% [58-64%]
Unweighted base	*3349*	*544*	*1647*	*1141*
Weighted base	*3307*	*555*	*1625*	*1108*

Base: All respondents at Wave 3.

Achieving higher than expected standard

For the Home Learning Environment Index a similar result to that seen for the expected standard was found for the higher than expected standard. A higher score on this index was associated with a greater likelihood of achieving the higher standard at KS2. For the

other home environment measures, however, a divergent pattern of results was found. The full findings for the higher expected standard can be found in Appendix 6: Complete results chapter regression tables: Table 51.

Discussion and conclusion

At KS1, a range of home environment factors, including the quality of the parent-child relationship, was found to have a significant influence on children's educational outcomes (Melhuish and Gardiner, 2021). At KS2, the association between home environment measures and attainment remains, but collectively, these measures only account for a small proportion of the overall variance in child attainment outcomes. This is true both when considering home environment measures from early childhood and when considering measures available around the time of the KS2 assessments, such as a more recent Home Learning Environment (HLE) index.

As at KS1, a better HLE index score in the early years is positively associated with a range of child attainment outcomes in primary school (ibid.). Therefore, children with a better HLE index score performed better in attainment assessments at both KS1 and KS2.

At KS2, a higher score on the authoritative parenting style was linked to a slightly higher likelihood of meeting the expected standard in reading, writing, and maths (combined). This association remained even after controlling for other home environment variables, socio-demographic factors, and childcare exposure.

The analysis also found that stricter parental limits on children's behaviour were somewhat associated with better outcomes in these subjects at KS2. This association persisted even when controlling for other home environment variables, socio-demographic factors, and childcare exposure. For comparison, at KS1 higher parental limit-setting was associated with a greater likelihood of children achieving the expected level in KS1 reading, maths, and science (ibid.).

However, some of the associations between the home environment and academic attainment at KS1 were not found to be repeated at KS2 for children's achievement of the expected standard in reading, writing and maths (combined). Higher permissive parenting was associated with poorer child outcomes on all KS1 measures (ibid.). Conversely, greater warmth in the parent-child relationship was associated with a higher likelihood of children achieving the expected level in KS1 reading, maths, and science. Higher levels of warmth were also linked to better performance in the phonics screening check. However, higher invasiveness in the parent-child relationship was associated with a lower likelihood of children achieving the expected level in KS1 Maths and Science. Comparable results for these measures of the home environment were not identified at KS2, once children's socio-demographics were controlled for.

This analysis of results at KS2 indicates that some associations between some home environment and attainment outcomes in the SEED study were less predictive over time. There is some evidence that the effects of ECEC on academic attainment may fade out over time, although this is not consistent (Melhuish et al., 2015) and other research has found effects lasting into adolescence (Taggart et al., 2015). Some studies have also found a pattern of fadeout followed by later reemergence of effects in other early childhood interventions (Bailey et al., 2023). Overall, the fact that certain individual early years home environment indicators are no longer statistically significant predictors of attainment at KS2 in this study should not distract from the bigger picture – children with a better home learning environment score are more likely to perform well academically also at KS2.

Chapter 4: Home environment and socio-emotional outcomes at Key Stage 2

Key Points

- Home environment factors during early years were associated with a child's socio-emotional outcomes at Key Stage 2 (KS2). The group of home environment factors explained 14.3% of the differences in total difficulties score and 9.7% of the differences in the prosocial behaviour at age 10 to 11.
- When considering the child's home environment also at KS2, a significant share of the variance in socioemotional outcomes is accounted for. Home environment factors during both early years and KS2 were able to explain 25.6% of the variance in the total difficulties score and 14.0% of the variance in the prosocial score.
- Higher levels of parental psychological distress when the children were younger, as well as higher values on the CHAOS, MORS invasiveness and parental limit scores during children's early years, were all found to be associated with a greater likelihood of children having a problematic total difficulties score at KS2.

Total difficulties and the home environment

The analysis explored the association between the children's overall socioemotional outcomes at KS2 and their home environment at different points in time. The main estimate of children's socio-emotional outcomes at KS2 comes from the Strengths and Difficulties Questionnaire (SDQ), with the total difficulties score as the main outcome. As in the previous chapter, the home environment measures include the Home Learning Environment (HLE) index, the two Mother Object Relations Scales (MORS) subscales (warmth and invasiveness), the Confusion, Hubbub, and Order Scale (CHAOS) scale, as well as measures of how far parents put limits on their children's behaviour, parental psychological distress, authoritarian parenting, authoritative parenting and permissive parenting. These are defined in 'Appendix 2: Variable descriptions'.

Table 17 presents an extract from the results. The association between home environment factors and the total difficulties score is explored, controlling for sociodemographic variables and ECEC exposure. It identifies both the household CHAOS scale and MORS invasiveness scale as associated with greater problems on the SDQ total difficulties score. It also found more parental limit setting to be associated with more difficulties on the total difficulty score.

Note that the child's SDQ total difficulties score at the time they attended ECEC is not included among the sociodemographic controls in this model, as children's early home environment may have influenced their earlier SDQ scores. As such, if a more disadvantaged home environment in children's early years caused them to have more socio-emotional problems in their early years, including this as another control in the model of outcomes at KS2 may confound the relationship between early years home environment and KS2 outcomes. Other model specifications are available in the Appendix 6: Complete results chapter regression tables: Table 52.

Table 17. Home environment and total difficulties score

	(1) **Total difficulties score (W7)**
How far parents put limits on their children's behaviour (average W1 and W2)	0.827*
CHAOS scale - average for W1, W2	0.219*
Parental mental distress (Waves 1 and 2)	-0.092
MORS invasiveness subscale (W2)	0.177***
MORS warmth subscale (W2)	-0.043
Authoritarian parenting score (W3)	0.956+
Authoritative parenting score (W3)	-0.086
Permissive parenting score (W3)	0.167
HLE index (average over W1, W2 and W3)	0.046

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. The predictors are correlated, so individual odds ratios should be interpreted with caution. See complete table in Appendix 6: Complete results chapter regression tables: Table 52.

Bivariate associations between the home environment and the total difficulties score

The home environment predictors included in these models were quite strongly correlated with each other, and this makes it difficult to attribute a unique effect to each measure. As such, rather than focussing on the coefficients while holding the other home environment factors constant, the reporting focusses on the bivariate relationship between the home environment measure and the SDQ total difficulties score.

The relationship between each individual home environment factor and the total difficulties score is explored through the descriptive statistics in the tables below, where the home environment variable was found to be a statistically significant predictor in the regression model. The full tables are provided in the separate tables annexe. In these tables, the association between individual home environment variables and the likelihood of having a problematic total difficulties score, defined as a score of 17 or more, is explored.

The CHAOS scale values recorded in the early years were positively associated with the total difficulties score at KS2. In other words, parents that reported issues such as struggling to stay on top of things or lacking calm in their home when the children were little were more likely to have children that experienced emotional difficulties at age 10 to 11. One in three children (34%) in families with high CHAOS scores during their early years had problematic total difficulties scores at KS2, compared with just one in ten (10%) children whose families had a low CHAOS scale assessment (Table 18).

Table 18: Whether had a problematic score on the SDQ total difficulties score at Wave 7, by household CHAOS score at Wave 3

Problems on the total difficulties score	Total	Low	Medium	High
No	83% [80-86%]	90% [86-92%]	81% [77-84%]	66% [55-75%]
Yes	17% [14-20%]	10% [8-14%]	19% [16-23%]	34% [25-45%]
Unweighted base	*2102*	*807*	*1171*	*124*
Weighted base	*2083*	*787*	*1172*	*124*

Base: All respondents at Wave 7.

Moderate parental psychological distress during early years, such as symptoms of depression or anxiety, was associated with children being more likely to experience social difficulties during KS2. Over one in three (36%) children to parents who reported a moderate level of distress when the child was very young went on to have problematic total difficulties scores at KS2, compared with one in eight (12%) children to parents who did not report such issues (Table 19).

Table 19: Whether had a problematic score on the SDQ total difficulties score at Wave 7, by level of parental psychological distress in early years

Problems on the total difficulties score	Total	No distress	Low Distress	Moderate distress	Moderately High Distress	High distress
No	83% [80-86%]	88% [86-90%]	77% [70-83%]	64% [54-72%]	*	*
Yes	17% [14-20%]	12% [10-14%]	23% [17-30%]	36% [28-46%]	*	*
Unweighted base	*2102*	*1522*	*321*	*192*	*35*	*
Weighted base	*2083*	*1473*	*335*	*202*	*40*	*

Base: All respondents at Wave 7.

The association between the MORS invasiveness subscale and the total difficulties was also positive. Children to mothers who had a high score on this scale, which measures issues such as a sense of unwelcome invasion or control by her infant in the early years, were more likely to experience socio-emotional difficulties at age 10 to 11. More than one in three (35%) children to mothers with a high MORS invasiveness score had a problematic total difficulties score at KS2, compared with one in eight (12%) children to mothers with a low MORS invasiveness score (

Table 20). A similar comparison for the warmth subscale was not possible, as this second MORS scale did not have enough mothers classified as reporting a low score.

Table 20: W7: SDQ total difficulties by MORS invasiveness subscale (banded - W2)

W7: SDQ total difficulties	Total	11 or less	12 - 17	17 - 35
No	83% [80-86%]	88% [86-90%]	76% [69-81%]	65% [57-72%]
Yes	17% [14-20%]	12% [10-14%]	24% [19-31%]	35% [28-43%]
Unweighted base	*2102*	*1417*	*486*	*172*
Weighted base	*2083*	*1361*	*479*	*204*

Base: All respondents at Wave 7.

Out of the variables capturing parenting style, the permissive parenting score has the most variation among the families in the SEED study. More than 100 of the children are assessed as having experienced low, medium and high levels of permissive parenting respectively. A permissive parenting style score during the early years, characterised by high responsiveness and low levels of psychological control, is associated with higher levels of total difficulties at KS2. Over one in four (28%) children who experienced high levels of permissive parenting went on to have problematic total difficulties scores at KS2, compared to one in eight (12%) children who experienced low levels of permissive parenting (Table 21).

Table 21: W7: SDQ total difficulties by Permissive parenting score (banded W3)

W7: SDQ total difficulties	Total	Low	Medium	High
No	83% [80-86%]	88% [85-90%]	80% [77-83%]	72% [63-79%]
Yes	17% [14-20%]	12% [10-15%]	20% [17-23%]	28% [21-37%]
Unweighted base	*2102*	*954*	*977*	*142*
Weighted base	*2083*	*947*	*949*	*154*

Base: All respondents at Wave 7.

The differences in socio-emotional wellbeing among children at KS2 by how much parents set limits on their behaviour were mixed. Children whose parents were more likely to set limits on their behaviour (score of 3 to 5) were more likely to experience difficulties on the SDQ total difficulties scale: 23% compared to 13% among children with a score of 2 to 3. On the other hand, in the group with very low parental limit setting

(score less than 2) 17% had a total difficulties score that identified problems with socio-emotional wellbeing. However, the confidence intervals for this percentage overlap with both the other groups and the base size for this group was small. The finding that higher parental limit setting was associated with higher socio-emotional problems at KS2 is also consistent with the results from the modelling.

Table 22 W7: SDQ total difficulties by Parental limit setting score (banded W2-3)

W7: SDQ total difficulties	Total	Less than two	Two or more but less than three	Three to five
No	83% [80-86%]	83% [77-88%]	87% [84-90%]	77% [73-80%]
Yes	17% [14-20%]	17% [12-23%]	13% [10-16%]	23% [20-27%]
Unweighted base	2102	325	1063	706
Weighted base	2083	334	1038	698

Base: All respondents at Wave 7.

Prosocial behaviour and the home environment

In addition to the subcomponents that make up the total difficulties score, prosocial behaviours are also captured by the Strengths and Difficulties Questionnaire (SDQ). The association between the prosocial score at KS2, capturing issues such as being considerate of people's feelings or sharing with others, and home environment variables is assessed in Table 23. Socio-demographic variables, excluding the total difficulties score during early years, and ECEC variables are included as controls. The overall fit of the model indicates that these variables combined can explain 20.9% of the variance in prosocial behaviour at KS2. The full model specification and alternative model specifications are displayed in Appendix 6: Complete results chapter regression tables: Table 53.

Table 23. Home environment and prosocial score

	(1) Prosocial score (W7)
How far parents put limits on their children's behaviour (average W1 and W2)	-0.162
CHAOS scale - average for W1, W2	-0.008
Parental mental distress (Waves 1 and 2)	0.008
MORS invasiveness subscale (W2)	-0.045**
MORS warmth subscale (W2)	0.051**
Authoritarian parenting score (W3)	0.072
Authoritative parenting score (W3)	0.215+
Permissive parenting score (W3)	-0.061
HLE index (average over W1, W2 and W3)	0.005

Note: The coefficients presented in the table are from Ordinary Least Squares regressions. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001. The predictors are correlated, so individual odds ratios should be interpretated with caution. See complete table in Appendix 6: Complete results chapter regression tables: Table 53.

Discussion and conclusion

Home environment factors during children's early years were found to be strongly associated with a child's socio-emotional outcomes at KS2. They were able to explain a greater share of children's socio-emotional outcomes than early years childcare experiences. This association was found for home environment factors both during the early years and when including data collected towards the end of primary school. When the more recent home environment measures (those collected at KS2) are considered, a better understanding of the variance in a child's socio-emotional outcomes was reached. That is to say, including both children's early years home life and current home life (at KS2) provided a better understanding of their socio-emotional outcomes at KS2. Further research may consider the relative importance of the home environment at different points in time, which this report did not address.

In particular, of the measures of the home environment included in the analysis, that the level of disruption and disorder in the household (the household CHAOS scale), the quality of the mother-child relationship (the MORS invasiveness scale) and how far parents set limits on children's behaviour were all found to be associated with greater socio-emotional difficulties. However, due to the strong interrelationship between these home environment factors (people with more difficulty in one area were also likely to experience problems in others) it is difficult in these models to attribute particular effects to each component of the home environment – and these findings should be treated with a degree of caution.

Concluding remarks

For KS2 attainment, the analysis established there is a small association between spending more time in formal ECEC and improved academic attainment at KS2, which is consistent with earlier findings at KS1. Similarly, some elements of the home environment such as the home learning environment, an authoritative parenting style and parental limits on children's behaviour also appear to be contributing factors for understanding which children succeed in reaching expected standards at KS2.

However, it is also important to reflect that any associations with a child's ECEC experience or home environment in the early years should be understood as a potential complementary factor to socio-demographic factors. Such circumstances, like whether a parent holds an academic qualification of Level 4 or above or if the child is in receipt of Special Educational Needs (SEN) support, are strongly associated with a child's likelihood of achieving the expected standard towards the end of primary school. For example, children from families at a financial disadvantage are less likely to achieve the expected standard at KS2. This KS2 analysis found that the positive association between increased formal ECEC use, and attainment is more pronounced for financially disadvantaged children, in line with the KS1 results, which indicates that formal ECEC is likely to be more important for these children's likelihood of reaching the expected standard throughout primary school.

For KS2 socio-emotional outcomes, this study lends clear further support to finding of earlier SEED impact analyses that the home environment and the quality of the parent-child relationship, both during the early years and towards the end of primary school, can have a clear, positive impact on children's outcomes. While again, factors that this analysis have classified as socio-demographic, including adverse events at an early age or receipt of SEN support, were also clearly linked to the children's socio-emotional outcomes at KS2, these results identify a notable additional contribution associated with the home environment.

References

Bailey D., Duncan G.J., Odgers C.L., Yu W. Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. J Res Educ Eff. 2017;10(1):7-39. doi: 10.1080/19345747.2016.1232459. Epub 2016 Nov 14. PMID: 29371909; PMCID: PMC5779101.

Berger, L.M., Panico, L., & Solaz, A. (2021). *The Impact of Center-Based Childcare Attendance on Early Child Development: Evidence From the French Elfe Cohort,* Demography, 58 (2), 419–450, https://doi.org/10.1215/00703370-8977274

Gutman, L.M., Joshi, H., Schoon, I. (2019). *Developmental Trajectories of Conduct Problems and Cumulative Risk from Early Childhood to Adolescence*, Journal of Youth and Adolescence, 48, 181-198, https://doi.org/10.1007/s10964-018-0971-x

Herbaut, E., Farges, G. & Giret, J. (2024). *Can early schooling at age 2 narrow the gaps in child development? Evidence from the French Elfe cohort*, Oxford Review of Education, https://doi.org/10.1080/03054985.2024.2305474

HM Treasury (2023) Spring Budget 2023, https://www.gov.uk/government/publications/spring-budget-2023

Kiernan, K. Crossman, S. & Phimister, A. (2024). *Families and Inequalities*, Dimensions of Inequality: The IFS Deaton Review Families, Oxford Open Economics, 3, 645-677, https://doi.org/10.1093/ooec/odad031

Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., Tawell, A., Slot, P., Broekhuizen, M., & Leseman, P. (2015). *A review of research on the effects of Early Childhood Education and Care (ECEC) upon child development*. CARE project; Curriculum and Quality Analysis and Impact Review of European Early Childhood Education and Care (ECEC).

Melhuish, E. (2016) 'Longitudinal research and early years policy development in the UK', *International Journal of Child Care and Education Policy* 10:3. https://doi.org/10.1186/s40723-016-0019-1

Melhuish, E., & Gardiner, J. (2017a). *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to Age Three: Research report,* Department for Education, London: DfE, https://assets.publishing.service.gov.uk/media/61964d19e90e0704478a9c75/SEED-Age_3_RESEARCH_REPORT.pdf

Melhuish, E., & Gardiner, J. (2017b). *Study of Early Education and Development (SEED): Study of Quality of Early Years Provision in England: Research report*, Department for Education, London: DfE, https://assets.publishing.service.gov.uk/media/5b438623e5274a376e24e115/Study_of_quality_of_early_years_provision_in_England.pdf

Melhuish, E., & Gardiner, J. (2018). *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age four years: Research report*, Department for Education, London: DfE, https://assets.publishing.service.gov.uk/media/61964e22e90e07043c35f311/SEED-Age_4_RESEARCH_REPORT.pdf

Melhuish, E., & Gardiner, J. (2020). *Study of Early Education and Development (SEED Impact study on early education use and child outcomes up to age 5 years: Research report*, Department for Education, London: DfE, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/867140/SEED_AGE_5_REPORT_FEB.pdf

Melhuish, E., & Gardiner, J. (2021). *Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age seven years: Research report,* Department for Education, London: DfE*,* https://assets.publishing.service.gov.uk/media/617a9eb2d3bf7f55fd843d0a/SEED_Age_7_Research_Brief.pdf

Melhuish, E., & Gardiner, J. (2023). *Equal Hours?: The impact of hours spent in early years provision on children's outcomes at age five, by socio-economic background*, The Sutton Trust, https://www.suttontrust.com/our-research/equal-hours/

Sammons, P., Sylva, K., Melhuish, E., Taggart, B., Elliot, K., & Siraj-Blatchford, I. (2002). *The Effective Provision of Pre-School Education (EPPE) Project: Measuring the impact of pre-school on children's cognitive progress over the pre-school period.* London: Institute of Education, University of London/ Department for Education and Skills.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., & Hunt, S. (2008). *Effective Pre-school and Primary Education 3-11 Project (EPPE 3-11): Influences on children's attainment and progress in Key Stage 2: Cognitive outcomes in Year 6*. London: Department for Children, Schools and Families

Speight, S., Bhatti, S., Keyes, A., and Parker, R. (2025) Study of Early Education and Development (SEED): Wave 7. Technical report. DfE Research Report.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart B. (2004). *The Effective Provision of Pre-School Education (EPPE) Project: Technical Paper 12: The Final Report: Effective Pre-School Education*, London, Department for Education and Skills

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart. B. (2012) *Effective Pre-school, Primary and Secondary Education 3-14 Project (EPPSE 3-14) - Final Report from the Key Stage 3 Phase: Influences on Students' Development form age 11-14*. Research Report DFE-RR202. London: Department for Education.

Taggart, B., Sylva, K., Melhuish, E., Sammons, P., & Siraj, I. (2015). Effective pre-school, primary and secondary education project (EPPSE 3-16+). UCL Institute of Education, University College London, Birkbeck, University of London, University of Oxford

Tamura, K., Morrison, J. & Pikhart, H. (2020). *Children's behavioural problems and its associations with socioeconomic position and early parenting environment: findings from the UK Millennium Cohort Study*, Epidemiology and Psychiatric Sciences, 29 (155), 1–14, https://doi.org/10.1017/S2045796020000700

Appendix 1: Technical details

Appendix 1 provides more details on the data and analysis methods used in this report. It describes the SEED study in more detail, the NPD data linkage conducted, the modelling approach used in this analysis, details of the SEED study's sampling and weighting approach, a more technical description of how to interpret the regression results, as well as a description of this analysis' limitations. It does not describe in detail the analysis variables used – these are dealt with in Appendix 2.

The Study of Early Education and Development (SEED)

The Study of Early Education and Development (SEED) is a major longitudinal study funded by the Department for Education (DfE) that has followed nearly 6,000 children from across England from the age of 2 years. The SEED study collects data on the type and amount of childcare that children received between the ages of 2 and 4, alongside factors such as demographics, the home learning environment, and the child's social and emotional wellbeing. In early waves, supplementary data on the quality of the childcare provision was also collected for a subset of the sample.

The original sample was designed to include a broad range of children from disadvantaged backgrounds who may be more eligible for funded early education. Therefore, each cohort of children within SEED was designed to have three equally sized subgroups:

- The 20% most disadvantaged families.
- The moderately disadvantaged (20-40%) families.
- The 60% least disadvantaged families.

For example, those from the 20% most disadvantaged families were oversampled so they made up one third of the total sample size. Allocation to each group was based on data from the Department of Work and Pensions and based on family's receipt of benefits and tax credits such as Jobseeker's allowance, Employment Support Allowance, and Child Tax Credit.

Data has been collected across several waves, starting when the children were aged 2 years and then collected annually while they were in Early Childhood Education and Care (ECEC), up to age 4. Further waves were then collected at age 5 years and during the COVID-19 pandemic, but no data from these waves is used in this report.[3] Most recently a seventh wave of data has been collected from children and their families from 2022 to 2023, when the children were in Year 6, or between 10 and 11 years old. At Wave 3, 3,930 parents completed the questionnaire, this reduced in Wave 7 to 2,146. This is a

[3] Waves 4, 5 and 6 were not included in the analysis for this report because the available sample size would be substantially restricted by including data from these waves, in particular with waves collected during the Coronavirus pandemic.

reduction of 1,784 productive cases, or 45% since Wave 3. The analysis in this report draws from survey responses from waves 1, 2, 3 and 7. Waves 1 – 3 were conducted in 2013 – 2016 (ages 2 to 4) and Waves 7 in 2022 and 2023 (ages 10 to 11).

National Pupil Database (NPD) linkage

In order to analyse the impact of ECEC and home environment factors on young people's Key Stage 2 attainment, survey data from SEED were linked with administrative pupil records held within the National Pupil Database (NPD). The NPD contains administrative data on all young people in England who engage with the state funded education system from Early Years education through to students in Key Stage 5. It includes information on their type of schooling and educational provision, educational demographics such as if the child has special education needs (SEN) or is in receipt of free school meals (FSM), and the young person's attainment.

For this analysis, data was requested from the NPD on the child's Key Stage 2 attainment, if the child was a 'Child in Need' (CIN) or a 'Looked after Child' (CLA), their gender and month of birth, and whether they were identified as having SEN or are in receipt of FSM. Linking was conducted by the Office for National Statistics (ONS) using probabilistic matching based on shared demographics held in both datasets. The data linking with NPD data resulted in an 85.2% match rate between the Wave 3 outcomes, increasing to 98.3% for the Wave 7 outcomes.

Data modelling

The primary models conducted for this analysis predict each attainment and SDQ outcome using 1) how much ECEC a child received in their early years, 2) the home environment factors, and 3) the socio-demographics of the child and their household. Follow up analysis for the role of ECEC was also conducted, adding two additional measures of ECEC: the age a child first received ECEC for more than 10 hours a week on average and the quality of the setting in which formal group ECEC was received.

In this analysis, children's academic attainment at KS2 is measured using a binary variable (either yes, the child has achieved the given academic standard, or no they have not), while the socio-emotional outcomes are measured using a continuous numeric score. As a result, two approaches were adopted to the modelling. For the binary attainment outcomes, binary logistic regression models were used. These models are suitable where the outcome of interest can take on only two values, yes or no. For the numeric socio-emotional scores, linear fixed effects models were used. These assume an approximately linear relationship between a predictor and an outcome. For example, that a person's total difficulties score on the SDQ will change by a similar amount if a child's average formal ECEC use changes from 5 to 10 hours per week as it would if it moved from 30 to 35 hours per week, when controlling for other variables.

For the analysis of academic attainment, the sample was taken from Wave 3 SEED respondents who were successfully matched to the NPD at KS2 (N = 3,930). This is because the sample is larger than respondents from Wave 7. The analysis of socio-emotional outcomes, however, relies on survey data collected at Wave 7 to provide the outcome measure. Therefore, the sample size was smaller (N = 2,146), limited to people who remained in the sample at Wave 7.

For the analysis of attainment, which has the larger sample size, in addition to the main models a version of the preferred model that stratifies it by the three disadvantage groups was produced. The SEED disadvantage groups split the sample into the 20% most disadvantaged families, the 20-40% most disadvantaged families, and the 60% least disadvantaged. Disadvantage is measured in determined based on the level of state financial support a family receives. Due to the oversampling of more disadvantaged families each of these groups make up approximately one third of the sample. The stratified model produced three separate sets of results, one for each disadvantage group, to allow for comparisons of the effects of different predictor and control measures on achieving the expected standard at KS2 for children who entered the SEED study with different levels of disadvantage.

Horvitz-Thompson heteroscedasticity robust standard errors were used in all models. The analysis was conducted using case-wise deletion, which excludes cases with any missing data on the variables used. This was deemed appropriate given the very low levels of missingness. Please refer to ‘Appendix 2: Variable descriptions’ for descriptive statistics on each variable.

Covariates

When socio-demographic control variables are included in a model these are:

- Ethnicity,
- Sex,
- SEED sample deprivation group,
- Experience of adverse events,
- Whether the child had a health, development, or behavioural issue,
- The parents’ highest qualification,
- The SDQ total difficulties score (Wave (W) 2-3),
- SEN status,
- FSM receipt and
- CIN status.

These were selected from a broader list of potential socio-demographic controls. Please refer to Appendix 2: Variable descriptions' and 'Appendix 3: Selection of socio-demographic control variables' for further information.

Sampling, weighting and clustering in the data

All analysis presented in this report was conducted using the relevant survey weights. The weights adjusted for non-response among participants to each wave of the SEED survey and the unequal sampling probabilities in the original sample (more disadvantaged children were over sampled).

The SEED sample is also clustered geographically by postcode district and by postcode sector within district. From each postcode district three sectors were sampled and from each postcode sector 5-6 families were sampled from each deprivation group (the sample was split into three deprivation groups). This means there is a level of geographical clustering in the data and the sample is not a truly random one.

This is the case with many face-to-face fieldwork surveys, where samples are clustered in order to increase the efficiency of fieldwork. To account for this clustering the primary sampling unit (PSU), the postcode district, was included in the complex survey design. This ensures that standard errors are calculated correctly – without this adjustment they would assume the data was randomly sampled, and the standard errors would be too small.

In the case of the SEED study there are two further potential elements of clustering in the data – school and ECEC provider. The sample was clustered by geography with the explicit intention of making it more likely that more than one participant would attend one ECEC provider, which would make the efficiency of fieldwork greater when assessing ECEC setting quality – a resource intensive step of the fieldwork. However, this also introduces an additional element of clustering which will only be partially accounted for by the PSU. As an ID variable for ECEC provider is not available in the SEED dataset, this cannot be accounted for in the analysis. In terms of clustering within schools, at waves 1 – 3 there were on average 1.4 children per school and 1.3 per school in waves 4 – 6. Given this relatively low level of clustering it was decided not to take account of this in the analysis.

For further detail of the survey methodology see the Study of Early Education and Development (SEED): Wave 7. Technical report (Speight et al., 2024).

Interpretation of results

Interpreting the binary logistic regression models

Having met the expected standards for different subjects is a binary, or dichotomous, variable. This means that a child can only have two outcomes: either they have met the expected standards, or they have not. The same applies to having met higher standards, the other core outcome analysed in this chapter. They way in which different variables (or predictors) influence a binary outcome variable is measured in this report with binary logistic regression models, also known as binary logit). In these models, the outcome of interest is transformed in a probability, allowing the exploration of how different predictors change the probability of experiencing the outcome (meeting expected standards), and how their relationship with the outcome changes when other variables are taken into account.

In this report, the estimates are presented as odds ratios. Odds ratios in categorical variables, such as being in receipt of Free School Meals (FSM), measure how likely someone with a specific characteristic (in this example, FSM) is to experience the outcome, compared to someone else with a defined and different characteristic, known as the reference category (non-FSM). For a continuous predictor, the odds ratio measures the change in likelihood that is associated with a one unit increase in the predictor; for example, in Table 1, the first model indicates that for each additional hour of formal group childcare, the odds of meeting the expected standards for different subjects increase 1.017 times. Odds ratios below one indicate a negative relationship and odds ratios above one indicates a positive relationship.

Each estimate is accompanied by a number of stars, ranging from zero to three. This indicates how confident we are that the association found in the data is statistically significant. Statistically significant means that it is unlikely to be the result of chance and can be confidently inferred to the population of interest in the study (if it were due to chance, the finding might only be specific to the people who took part in the survey). Zero stars indicates that the association is not statistically significant, as the probability that it occurred by chance is greater than 5% ($p > 0.05$). One star (*) indicates less than a 5% probability of the observed relationship having occurred by chance ($p < 0.05$). This decreases to less than a 1% probability ($p < 0.01$) for two stars (**) and less than a 0.1% probability ($p < 0.001$) for three stars (***).

In Appendix 6: Complete results chapter regression tables the full regression tables are shown which include a number of additional statistics. These include the standard error below each coefficient in square brackets, reflecting the variability or uncertainty around that estimate. A smaller standard error indicates a more precise estimate. To summarise the goodness-of-fit of the different models (how well each model fits the data, and can explain the variation in the outcome), the Cox-Snell Pseudo-R-squared is also shown. This parameter can be used to compare different models where they were run on the

same sample and can help understand which models are better at predicting the outcome of interest (the higher the value, the better the model).

Effect size: Overall attainment and childcare hours

The modelled results indicate that each hour of additional weekly attendance in day nurseries, nursery classes or schools and playgroups between ages 2 and 4 was associated with an 1.5% increase in the odds (odds ratio 1.015) of achieving the expected standard in reading, writing and maths (combined) at the end of primary school, con-trolling for socio-demographic factors, home environment factors, and the amount of other forms of childcare received. Ten additional hours of formal group childcare per week was associated with an 16% increase (exp(10 * ln(1.015)) = 1.16) in the odds of achieving the expected standard. This corresponds to an increased probability of achieving the expected standard of about 3 percentage points.

The terms ‘odds’ and ‘probability’ are often used interchangeably in everyday conversations. In statistics, these terms are related but calculated in different ways. Probability describes the likelihood that an event, such as a child achieving the expected standard, will occur. Odds is the ratio between the probability that the event will occur and the probability that it will not occur. If the difference between these two concepts is missed, the effect size can easily be misinterpreted.

Let us consider a child selected at random. Assuming they have a 61% probability of achieving the expected standard, as this is the population weighted average in the study. The corresponding odds is 1.564 (0.61/(1-0.61) = 1.564). For this child, the model indicates that their odds of achieving the expected standard would have in-creased to 1.814 if they had attended 10 additional hours of formal group childcare provision (a 16% improvement of the original odds: 1.16 * 1.564 = 1.814). That is the same as a probability of 64% (1.814/(1+1.814)=0.64).

The probability of achieving the expected standard hence increases from 61% to 64%. In other words, the increased probability of achieving the expected standard associated with a 10 hour increase in formal group childcare is about 3 percentage points, controlling for socio-demographic factors, home environment factors, and the amount of other forms of childcare received.

Interpreting the OLS regression models

SDQ scales can be considered as continuous variables meaning that the variations in the scale’s scores can be analysed using linear regression models, also known as Ordinary Least Squares regressions. This statistical model estimates to what extent changes in the SDQ scale scores are influenced by a range of different variables, or predictors.

For continuous predictors, the coefficient represents the expected change in the outcome for each one-unit increase in that predictor, holding all other variables constant. For

example, model 1 in Table 6 indicates that, when controlling for other variables, SDQ increases by 0.046 for each additional hour of formal individual childcare.

For categorical predictors, one category of the variable is selected as the reference group (for example, male), while a coefficient is assigned to the other groups (for example, female). This coefficient estimates the mean difference in the outcome variable that can be identified between a given group and the reference category.

Statistical significance are presented in the same way as for the logistic regression models.

As with the logistic regression models, in Appendix 6: Complete results chapter regression tables the full regression tables are shown which include a number of additional statistics. The goodness-of-fit of the model to the data and the predictive ability of the model are summarised in the R-squared parameter. The R-squared value ranges from 0 to 1, with a higher value indicating that the model can explain more of the variability in the outcome. These parameters can also be interpreted in absolute terms: a R-squared value of 0.41 means that the model can explain 41% of the total variance of the outcome variable. The standard errors are presented below each coefficient in square brackets.

Effect size: Socio-emotional wellbeing outcomes and childcare hours

The modelled results indicate that each hour of additional weekly attendance in day nurseries, nursery classes or schools and playgroups between ages 2 and 4 was associated with an 0.032 increase in a child's total difficulties score, once their socio-demographic and home environment characteristics are controlled for, as well as the number of hours spent in other types of childcare.

Thus, ten additional hours of formal group childcare per week then would be associated with a 0.32 increase in total difficulties score. To put this in perspective, a child's potential values on the total difficulties score range from 0-34 and the average score among children at Wave 7 of the survey was 9.7, with a standard deviation of 6.7.

Limitations

Firstly, it should be noted that childcare use is almost universal in this study, as it is among children today. Only 19 children in the sample (0.5%) did not attend any ECEC before starting school. Therefore, these results should be interpreted as the effects associated with variations in the amount and type of childcare attended. They cannot be read as the impact of attending or not attending childcare.

Secondly, this study is not randomised. As with other cohort studies, the primary limitation of this approach to data collection is the lack of control over the choices families have made. Children have not been randomly assigned to receive different levels or

types of childcare, nor have some been assigned to a control group receiving no childcare to provide a baseline for comparison with. As a result, while the SEED data provides rich and detailed insights by allowing researchers to follow thousands of families over time, it also has inherent limitations. Therefore, when interpreting these results, it is important to consider that confounding factors in a child's or family's circumstances may be linked to both factors of interest, such as higher childcare attendance in the early years and the child's later mental wellbeing or school attainment.

The KS1 report also noted that Special Educational Needs (SEN) could be a confounding variable in the analysis, as children with SEN may be less likely to use formal ECEC and are likely to have poorer cognitive and educational outcomes (Melhuish and Gardiner, 2021). This KS2 report controls for SEN status at the age of 10 to 11 and still finds a small association between previous childcare attendance in formal group settings and KS2 attainment in reading, writing, and maths (combined). This finding provides some support for the view that SEN is not a confounding variable in this relationship.

Finally, though the SEED study contains rich data on family circumstances, some family characteristics such as more detailed aspects of parents' professions are not covered, so this analysis has not controlled for such potentially confounding factors.

Appendix 2: Variable descriptions

Outcomes

Attainment

Achieved Expected Standard

A set of binary outcomes demonstrating if a child was working at expected standard at Key Stage 2 were produced across key subjects. The expected standard was indicated by a standardised scaled score of at least 100 in the relevant assessment across Maths, Reading, and Grammar, Punction, and Spelling (GPS). The standardised scores were defined as the raw score that the pupil achieved on a given assessment, adjusted for the particular difficulty of the questions in the assessment for a given year. Achieving the expected standard in Writing and in science was determined by a teacher assessment instead. A combined metric was also produced indicating if a young person had achieved expected standard in Reading, Writing, and Maths, this outcome serves as the primary outcome for the report.

Table 24. Distribution of KS2 attainment outcomes: expected standard

Outcome	Value	Weighted Proportion	Unweighted Base
KS2 Grammar, Punctuation, and Spelling: pupil met expected standard	No	25.88%	891
KS2 Grammar, Punctuation, and Spelling: pupil met expected standard	Yes	74.12%	2458
KS2 Maths: pupil met expected standard	No	26.06%	885
KS2 Maths: pupil met expected standard	Yes	73.94%	2464
KS2 Reading: pupil met expected standard	No	24.25%	818
KS2 Reading: pupil met expected standard	Yes	75.75%	2531

Outcome	Value	Weighted Proportion	Unweighted Base
KS2 Writing: pupil met expected standard (teacher assessed)	No	27.11%	920
KS2 Writing: pupil met expected standard (teacher assessed)	Yes	72.89%	2429
KS2 Science: pupil met expected standard (teacher assessed)	No	18.20%	616
KS2 Science: pupil met expected standard (teacher assessed)	Yes	81.80%	2733
KS2 Reading, Writing, and Maths: pupil met expected standard	No	39.27%	1327
KS2 Reading, Writing, and Maths: pupil met expected standard	Yes	60.73%	2022

Achieved Higher Standard

An additional set of binary outcomes were included in the analysis showing if a young person was working at greater than expected standard in Maths, Reading, Writing, or GPS. A high score in Match, Reading, or GPS was indicated by a scaled score of greater than 110, as writing is teacher assessed, achieving a higher standard is instead defined as ‘working at greater depth’. A combined metric was also produced indicating if a young person had achieved higher than expected standard in Reading, Writing, and Maths, this outcome serves as the secondary outcome for the report.

Table 25. Distribution of KS2 attainment outcomes: greater than expected standard

Outcome	Value	Weighted Proportion	Unweighted Base
KS2 Grammar, Punctuation, and Spelling: pupil met greater than expected standard	No	70.47%	2394
KS2 Grammar, Punctuation, and Spelling: pupil met greater than expected standard	Yes	29.53%	955

Outcome	Value	Weighted Proportion	Unweighted Base
KS2 Maths: pupil met greater than expected standard	No	76.61%	2579
KS2 Maths: pupil met greater than expected standard	Yes	23.39%	770
KS2 Reading: pupil met greater than expected standard	No	69.39%	2332
KS2 Reading: pupil met greater than expected standard	Yes	30.61%	1017
KS2 Writing: pupil met greater than expected standard (teacher assessed)	No	86.86%	2920
KS2 Writing: pupil met greater than expected standard (teacher assessed)	Yes	13.14%	429
KS2 Reading, Writing, and Maths: pupil met greater than expected standard	No	92.33%	3099
KS2 Reading, Writing, and Maths: pupil met greater than expected standard	Yes	7.67%	250

Social and emotional wellbeing

To measure the young people's social and emotional wellbeing, the parents completed the Strengths and Difficulties Questionnaire (SDQ) about their child. The questionnaire asks a set of 25 questions, such as if the child is social with other children or if they are nervous in unfamiliar situations. Different items are then used to derive a set of five sub-scales and an overall difficulties score, each showing if the child's behaviours can be regarded as problematic or not. It is important to note that the threshold scores for problematic behaviours change depending on the age of the child and who completed the questionnaire.

Total difficulties score

The total difficulties score measures the overall social and emotional wellbeing of the child and is being used as the primary outcome for measuring children's wellbeing in this report. It is based on combined score of four of the SDQ sub-scales, excluding the child's prosocial score. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of 17 or greater was defined as a problematic score.

Prosocial score

The prosocial subscale indicates a child who is considerate of other's feelings, readily shares with others and demonstrates kindness to others. Prosocial scores are reverse coded, that is a lower score indicates a more problematic behaviour. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of four or lower indicates a problematic behaviour. As the Pro-Social subscale is not used to derive the total difficulties score, it will act as a secondary outcome of interest for this report.

SDQ subscales

The Emotional Problems subscale assess if a child is often unhappy or downhearted, may seem nervous or clingy, or if they seem to worry often. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of five or higher is defined as a problematic score. The Hyperactivity subscale assess if a child is often restless and overactive, indicating a child may be easily distracted. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of seven or higher is defined as a problematic score. The Peer Problems subscale measures if a child is often solitary and tends to play alone, it may also indicate children who are picked on or bullied by other children. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of four or higher is defined as a problematic score. The Conduct Problems subscale shows if a child often experiences tantrums, may fight with other children or shows bullying behaviour, it also captures children who are often dishonest. For a child aged 10 to 11 and the questionnaire completed by the parent, a score of four or higher is defined as a problematic score.

Table 131. Distributions of SDQ outcomes

Outcome	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
SDQ conduct problems scale	1.53	1	1.69	2109
SDQ emotional problems scale	2.71	2	2.40	2111

Outcome	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
SDQ hyper-activity problems scale	3.71	3	2.76	2107
SDQ peer problems scale	1.78	1	1.89	2108
SDQ Prosocial scale	8.47	9	1.83	2110
SDQ total difficulties scale	9.71	8	6.74	2102

Predictors

Early childhood education and care

Hours of ECEC attended

To measure the volume of ECEC provision a child attended, data was collected on the number of hours they attended per week at different types of childcare provision. Childcare type grouped the settings in which ECEC is provided by whether they are a formal setting that is eligible for government funding (e.g. day nurseries, nursery classes, nursery schools and playgroups), or informal care (e.g. from relatives, friends, nannies or neighbours). It also split the formal care by whether it is delivered by an individual, such as a childminder, or by a group of childcare providers. Resulting in three continuous measures of the average number of hours a child attended each of the three settings a week, aged 0 to 4. A categorical measure of childcare hours was also produced be separating the continuous scores into quartiles.

ECEC setting quality measure

Sustained Shared Thinking and Emotional Well-being (SSTEW) scale

To assess the quality of the formal childcare some children were receiving, 1000 formal childcare providers, that children within the SEED study attended, were assessed. The Sustained Shared Thinking and Emotional Well-being scale is a measure of staff child interaction quality, the study uses an average of the scores collected at wave one and two. An additional derived variable has been produced, the first recoded the scale as a categorical measure based in the quartiles of the score and adds a fifth category that includes those who did not have their setting assessed, this is to allow for the full sample to be included in the regression models.

Table 26. Distributions of ECEC predictors

Predictor	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
SSTEW (average over W1 and W2)	4.67	4.77	1.16	1465
Hours of Informal Childcare Ages 0 to 4 (average over W1 - W3)	4.32	0.20	7.16	3930
Hours of formal group childcare ages 0 to 4 (average over W1 - W3)	9.89	7.20	7.53	3930
Hours of formal individual childcare ages 0 to 4 (average over W1 - W3)	1.42	0.00	4.66	3930

Home environment

Parental limit setting

To measure how far parents put limits on their children's behaviour, the total number of limitations that parents put on their child's behaviour was averaged across waves 1 and 2, a categorical variable was then derived using this continuous score. A categorical measure of this variable was also produced grouping together those parents that put in place less than two, two to three, or three to five limits.

CHAOS scale

Disorder in the household was measured as the average score on the CHOAS scale across the first two waves of data collection, the scale collects data on experiences such as. The Confusion, Hubbub, and Order Scale measures the level of confusion and disorganisation that is experienced within the child's home environment with a higher score indicating a greater level of CHAOS in the home. A categorical measure of household disorder was also produced with the following three categories:

- A score of 0 – 7: Low.
- A score of 7 to 11: Medium.
- A score of greater than 11: High.

Parental mental distress

To measure the level of psychological distress in parents, an average of the Kessler 6 psychological distress score was taken across the first two waves. The data was reported from whichever parent completed the survey and is indicative of symptoms of mental health issues like depression or anxiety. The Kessler 6 is measured on a scale with a higher score indicating a higher level of psychological distress factors reported by the parent. A categorical measure of parental psychological distress was also produced with the follow categories:

- No distress – score of 0.
- Low distress – a score of 0 to 0.5.
- Moderate distress – a score of between 0.5 to 1.5.
- Moderately high distress – a score of between 1.5 to 2.5.
- High distress - a score of 2.5 or greater.

Mother object relations scales (MORS)

To measure the relationship between the parent and their child, the Mother Object Relations Scales (MORS) were asked in the second wave of data collection and were used to produce two separatee measures. The first is a subscale for the warmth of the parental relationship, measuring aspects such as affection, or how much they do things together. The second subscale measured parental invasiveness, that is a measure of the level of conflict in their relationship, drawing on elements like feeling annoyed with the child, or viewing the child as demanding attention. Both MORS are measured on a scale with a higher score indicating greater warmth or invasiveness on the respective scale.

Parenting scores

To determine the parent styles used when the young people were aged 3, the survey asked several questions about parenting behaviours such as discipline within the home, if they consider their child their friend, and how much freedom they let their child have. These items were then used to derive scales for three distinct styles of parenting:

- An authoritative parenting style score – characterised by high responsiveness and high levels of psychological control.
- An authoritarian parenting style score – low responsiveness and high levels of psychological control.
- A permissive parenting style score – by high responsiveness and low levels of psychological control.

These form three scales with greater scores across all scales indicating a strong level of the relevant parenting style. Categorical variables for each score were also produced with the following categories:

- Low – a score of less than two.
- Medium – a score of two to three.
- High – a score of three to five.

Home Learning Environment (HLE) index

The HLE is an average measure across the first three waves that considers activities in the home that promote and provide learning opportunities for the child; such as a parent reading with their child, parent taking child to library, painting/drawing, playing with letters/numbers, songs/rhymes etc. The index is measured on a scale with a higher value indicating a better home learning environment. A categorical measure of HLE was also produced by separating the score into quintiles.

Table 27. Distributions of home environment predictors

Predictor	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
How far parents put limits on their children's behaviour (average W1 & W2)	2.65	2.64	0.65	3899
CHAOS scale (average for W1 & W2)	7.95	8.00	2.00	3930
MORS invasiveness subscale (W2)	9.85	9.00	5.01	3846
MORS warmth subscale (W2)	31.51	32.00	3.23	3854
Authoritarian parenting score (W3)	1.64	1.58	0.40	3850
Authoritative parenting score (W3)	4.16	4.20	0.47	3844
Permissive parenting score (W3)	2.03	2.00	0.58	3854
HLE index (average over W1, W2, & W3)	24.81	25.33	6.23	3930

Socio-demographic controls

The following list of socio-demographic characteristics were controlled for in a selection of the models.

Ethnicity

A categorical measure of the child's ethnicity was used, containing the groups Asian, Black, Mixed, and Other included, with White being used as the reference group in the regression analysis.

Sex

A binary measure of the child's sex was included, with 'Male' being used as the reference category for the regression analysis.

Disadvantage group

To maximise the study's ability to make comparisons of child outcomes across the spectrum of eligibility for funded early education for two-year-olds, each cohort of children within SEED was designed to have three subgroups:

(1) The 20% most disadvantaged families.

(2) Moderately disadvantaged families (20-40%).

(3) The 60% least disadvantaged families.

The three subgroups were sampled in equal proportion, i.e., such that each group made up around a third of the sample in each cohort. As the three groups were not of equal size in the population, a weighted sampling approach was used to create as close to an equal probability sample as possible. The selection weights were equal to the ratio of the desired proportion (one third) to the population proportion in each cohort.

Families were put into groups by Department for Work and Pensions (DWP) prior to sampling using the following criteria:

- The *20% most disadvantaged families* had a parent in receipt of one of the following benefits or tax credits:[4]
    - Income-based Jobseeker's Allowance (JSA-IB).
    - Income-related Employment Support Allowance (ESA-IR).
    - Income Support (IS).

[4] The full DfE eligibility criteria from September 2013 were: (I) All two-year-olds who were looked after by their local authority; (ii) two-year-olds whose family received one of the following were also eligible: income support; income-based Jobseeker's Allowance (JSA); income-related Employment and Support Allowance (ESA); support through part 6 of the Immigration and Asylum Act; the guaranteed element of State Pension Credit; Child Tax Credit (but not Working Tax Credit) and had an annual income not over £16,190; the Working Tax Credit 4-week run on (the payment you get when you stop qualifying for Working Tax Credit) or Universal Credit.

    - Guaranteed element of the State Pension Credit (PC with Guarantee Credit).
    - Child Tax Credit *only* (not in receipt of an accompanying Working Tax Credit award) with household gross earnings of less than £16,190.
- The *moderately disadvantaged group (20-40%)* had a parent in receipt of Working Tax Credits with household gross earnings of less than £16,190.[5]
- The *60% least disadvantaged group* had parents not in receipt of any of the qualifying benefits or tax credits.

For the regression analysis, children from the 20% most disadvantaged families were used as the references group.

Adverse childhood events

A continuous measure of the number of adverse events experienced by the child, such as a death in the family or parental divorce, was included to control for the impact of such experiences on the educational and socio-emotional outcomes of the child.

Health, development, or behavioural Issue

A binary measure showing if the child has a been identified as having a health, development, or behavioural issue between the ages of 0 and 4. Those who were identified as having such an issue were used as the reference group in the regression analysis.

Parental qualifications

A categorical measure was derived that showed the highest qualification, either academic or vocational, within the household from level 1 through to level 8. Parents who did not hold any qualifications were used as the reference group in the regression analysis.

Early years socio-emotional wellbeing

To control for the child's socio-emotional wellbeing during early years, an average of the child's total difficulties score during wave 2 and wave 3, when the child was aged 3 or 4.

[5] From September 2014, the eligibility criteria included two-year-olds who met any one of the following criteria: eligibility criteria also used for free school meals; if their families received Working Tax credits and had annual gross earnings of no more than £16,190 per year; if they had a current statement of special educational needs (SEN) or an education, health and care plan; if they attracted Disability Living Allowance; if they were looked after by their local authority; or if they had left care through special guardianship or through an adoption or child arrangements order.

Do avoid issues with multicollinearity, this control variable is not included in the regression models used to explore children's socio-emotional outcomes at KS2.

Special educational needs

A categorical variable was derived using information form the National Pupil Database (NPD) to show if the child was identified as having a Special Education Need (SEN) such as a sensory disability or autism. An additional category was produced showing if the child was identified as having SEN and they were supported by an Education Health and Care Plan (EHCP). Those children who were not identified as having SEN and were not supported by an EHCP acted as the reference group for the regression analysis.

Receipt of free school meal

To further understand the impact of some children being from disadvantaged background, a binary measure was included, taken from the NPD, in the demographic controls to show if a child was in receipt of free school meals (FSM). Those children who were not eligible were used as the reference group for the analysis.

Child in need status

A binary variable was produced with data taken from the NPD that showed if a child was classed as a child in need. A child in need includes all disabled children as well as those who need additional support and protection because of risks to their development. Looked after children are also included in the umbrella term of children in need, but as there are so few looked after children in the analysis, and they are a vulnerable group, they could not be identified as a separate group due to the disclosive risk. Children not identified as a child in need were used as the references category for the regression analysis.

Table 28. Distributions of socio-demographic controls

Socio-demographic control	Value	Proportion (Weighted)	Base (Unweighted)
Ethnicity of child	White	80.50	3299
Ethnicity of child	Asian	7.75	243
Ethnicity of child	Black	5.30	156
Ethnicity of child	Mixed	5.79	208
Ethnicity of child	Other	*	*

Socio-demographic control	Value	Proportion (Weighted)	Base (Unweighted)
Sex	Male	52	1116
Sex	Female	48	1030
Deprivation group	20% most deprived	22.42	958
Deprivation group	20%-40% most deprived	34.17	1398
Deprivation group	60% least deprived	43.41	1574
Child had a health, development, or behavioural issue (W1 – W7)	Yes	14.54	603
Child had a health, development, or behavioural issue (W1 – W7)	No	85.46	3327
Highest Qualification in Household, academic or vocational (W1 – W7)	No Qualifications Reported	5.83	74
Highest Qualification in Household, academic or vocational (W1 – W7)	Level 1	3.96	73
Highest Qualification in Household, academic or vocational (W1 – W7)	Level 2	20.22	371
Highest Qualification in Household, academic or vocational (W1 – W7)	Level 3	19.30	404
Highest Qualification in Household, academic or vocational (W1 – W7)	Level 4 or above	47.60	1181

Socio-demographic control	Value	Proportion (Weighted)	Base (Unweighted)
Highest Qualification in Household, academic or vocational (W1 – W7)	Other	3.09	37
Pupil is in receipt of Special Educational Needs (SEN) provision	No special educational need	79.72	2662
Pupil is in receipt of Special Educational Needs (SEN) provision	SEN support	16.21	536
Pupil is in receipt of Special Educational Needs (SEN) provision	EHCP	4.07	147
Pupil has been in receipt of Free School Meal (FSM) provision within the last 6 years	No	68.99	2315
Pupil has been in receipt of Free School Meal (FSM) provision within the last 6 years	Yes	31.01	1030
Pupil classed as a Child in Need (CIN) at time of data collection	No	97.38	3251
Pupil classed as a Child in Need (CIN) at time of data collection	Yes	2.62	98

Socio-demographic control	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
Average number of adverse events in the past 12 months (W1 and W2)	0.32	0.00	0.41	3930

Socio-demographic control	Weighted Mean	Weighted Median	Weighted SD	Unweighted Base
SDQ total difficulties score (W2 and W3)	8.97	8.00	4.71	3913

Appendix 3: Selection of socio-demographic control variables

Nine sociodemographic variables are included as control variables in regression models where 'Controls: Socio-demographic' are marked 'Yes'. These are:

- Ethnicity of child (White / Asian / Black / Mixed / Other).
- Sex (Male / Female).
- Deprivation group (20% most deprived, 20%-40% most deprived, 60% least deprived).
- Adverse events: Child had a health, development, or behavioural issue recorded in SEED survey waves 1, 2 or 3 (Yes / No).
- Highest parental qualification (No Qualifications Reported / Level 1 / Level 2 / Level 3 / Level 4 or above / Other).
- The child's SDQ total difficulties score at SEED survey waves 2 and 3 (arithmetic average).
- Special Educational Needs (SEN) status (No special educational need / SEN support / EHCP).
- Free School Meals (FSM) receipt (Yes / No).
- Children In Need (CIN) status (Yes / No).

Twenty one demographic controls were considered for inclusion. The full list of controls was considered too large to be included in the final models, as it would potentially lead to an overfitting of the models to this dataset. To assess which to include in this modelling at KS2 preliminary model was produced of the two key outcomes of interest. The two outcomes used was whether a child achieved the expected standard in reading, writing and maths (combined) and whether the total difficulties score indicated a child had a mental health problem.

The variables that provided additional explanatory power in one of these models were included as one of the demographic control variables in the analysis. If a variable was found to be statistically significant at the 10% level for either SDQ or attainment outcomes this was included in the final list of controls for both attainment and SDQ outcomes to ensure consistency in the analysis.

The results of the preliminary modelling, including only the socio-demographic covariates are presented below. These were modelled using binary logistic regression. The table below presents an overall measure of whether each variable contributed to the fit of the model, after controlling for all of the other socio-demographic characteristics.

Table 29. Socio-demographic control model

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
Intercept	0.844	1.728	0.054+	0.013***
standard error	[0.553]	[1.080]	[0.083]	[0.011]
Child's Birth Weight in KGs	1.096	-	1.084	-
standard error	[0.073]	-	[0.146]	-
Child's age in months	1.013	-	0.948	-
standard error	[0.012]	-	[0.028]	-
Mothers age at birth of cohort child: 20 to 29	0.854	-	1.493	-
standard error	[0.196]	-	[0.751]	-
Mothers age at birth of cohort child: 30 to 39	0.972	-	1.038	-
standard error	[0.231]	-	[0.553]	-
Mothers age at birth of cohort child: 40 or older	0.867	-	0.51	-
standard error	[0.279]	-	[0.351]	-
Ethnicity of child: Asian	1.906+	1.039	0.541	0.438+
standard error	[0.537]	[0.423]	[0.280]	[0.214]
Ethnicity of child: Black	1.337	1.502	1.547	1.423
standard error	[0.299]	[0.433]	[0.802]	[0.725]
Ethnicity of child: Mixed	1.198	1.057	1.361	1.335
standard error	[0.237]	[0.254]	[0.509]	[0.459]
Ethnicity of child: Other	2.324	2.145	1.035	0.307

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
standard error	[1.973]	[1.661]	[1.526]	[0.232]
Sex: Female	1.249*	1.135	1.14	1.112
standard error	[0.117]	[0.113]	[0.215]	[0.198]
Age (Years)	1.018	-	1.011	-
standard error	[0.011]	-	[0.027]	-
Deprivation group: 20%-40% most deprived	0.896	1.122	0.649	0.541*
standard error	[0.156]	[0.257]	[0.209]	[0.126]
Deprivation group: 60% least deprived	1.041	1.579+	0.632	0.521*
standard error	[0.192]	[0.381]	[0.222]	[0.136]
Average number of adverse events in the past 12 months (W1 and W2)	0.981	1.009	1.583+	1.626*
standard error	[0.114]	[0.160]	[0.334]	[0.324]
Household Income Banded (Wave 1 and 2): £10,000 - £19,999	0.97	-	1.019	-
standard error	[0.140]	-	[0.302]	-
Household Income Banded (Wave 1 and 2): £20,000 - £29,000	0.993	-	1.079	-
standard error	[0.183]	-	[0.404]	-

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
Household Income Banded (Wave 1 and 2): £30,000 - £49,000	0.992	-	1.231	-
standard error	[0.186]	-	[0.531]	-
Household Income Banded (Wave 1 and 2): £50,000 or more	1.327	-	0.628	-
standard error	[0.312]	-	[0.360]	-
Highest parental socio-economic status (W1-W2): Lower managerial	0.847	-	0.992	-
standard error	[0.124]	-	[0.319]	-
Highest parental socio-economic status (W1-W2): Intermediate occupations	0.794	-	0.847	-
standard error	[0.141]	-	[0.353]	-
Highest parental socio-economic status (W1-W2): Small employers/own account workers	0.828	-	0.326+	-
standard error	[0.218]	-	[0.175]	-
Highest parental socio-economic status (W1-W2): Lower supervisory	0.795	-	0.798	-
standard error	[0.208]	-	[0.410]	-

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
Highest parental socio-economic status (W1-W2): Semi-routine	0.524*	-	0.549	-
standard error	[0.111]	-	[0.224]	-
Highest parental socio-economic status (W1-W2): Routine	0.504*	-	1.039	-
standard error	[0.138]	-	[0.475]	-
Highest parental socio-economic status (W1-W2): Not working	0.794	-	0.432	-
standard error	[0.393]	-	[0.366]	-
Child had a health, development, or behavioural issue (W1 - W3): No	1.133	0.97	1.516	1.472
standard error	[0.181]	[0.192]	[0.487]	[0.422]
Ever lived in a lone parent household (W1 - W3): Respondent living with a partner	0.832	-	0.845	-
standard error	[0.143]	-	[0.211]	-
IMD quintiles (latest value W1-W3): 2	1.252	-	0.81	-
standard error	[0.174]	-	[0.231]	-
IMD quintiles (latest value W1-W3): 3	1.088	-	0.922	-
standard error	[0.166]	-	[0.263]	-

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
IMD quintiles (latest value W1-W3): 4	1.349+	-	0.922	-
standard error	[0.213]	-	[0.281]	-
IMD quintiles (latest value W1-W3): Most deprived	1.311	-	1.08	-
standard error	[0.225]	-	[0.300]	-
Household's accommodation tenure (latest value W1-W3): Social renting	0.781	-	1.061	-
standard error	[0.110]	-	[0.286]	-
Household's accommodation tenure (latest value W1-W3): Renting from a private landlord	0.691*	-	1.026	-
standard error	[0.091]	-	[0.270]	-
Household's accommodation tenure (latest value W1-W3): Living here rent free (e.g. with relatives or friends	0.811	-	0.52	-
standard error	[0.238]	-	[0.309]	-
Ever lived in a workless household (waves 1, 2 and 3): Has lived in a workless household	-	-	-	-
Ever lived in a workless household (waves 1, 2	1.249	-	0.781	-

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
and 3): Not lived in a workless household				
standard error	[0.265]	-	[0.213]	-
Number of siblings (latest value W1-W3): One	1.188	-	0.885	-
standard error	[0.147]	-	[0.204]	-
Number of siblings (latest value W1-W3): Two	0.958	-	0.952	-
standard error	[0.121]	-	[0.294]	-
Number of siblings (latest value W1-W3): Three	0.807	-	0.936	-
standard error	[0.158]	-	[0.371]	-
Number of siblings (latest value W1-W3): Four	0.618	-	0.406	-
standard error	[0.253]	-	[0.301]	-
Number of siblings (latest value W1-W3): Five or more	0.726	-	0.169	-
standard error	[0.275]	-	[0.228]	-
w1 to w7: Highest Qualification (academic or vocational) in Household: Level 1	1.391	1.024	0.802	0.993
standard error	[0.454]	[0.517]	[0.380]	[0.527]

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
w1 to w7: Highest Qualification (academic or vocational) in Household: Level 2	1.154	1.202	0.869	0.844
standard error	[0.292]	[0.492]	[0.439]	[0.379]
w1 to w7: Highest Qualification (academic or vocational) in Household: Level 3	1.609	1.488	1.018	1.064
standard error	[0.452]	[0.644]	[0.530]	[0.510]
w1 to w7: Highest Qualification (academic or vocational) in Household: Level 4 or above	2.057*	2.279+	0.589	0.629
standard error	[0.528]	[0.997]	[0.278]	[0.288]
w1 to w7: Highest Qualification (academic or vocational) in Household: Other	2.377+	1.542	2.734	1.389
standard error	[1.064]	[0.932]	[2.106]	[1.081]
SDQ total difficulties score (W2 and W3)	0.976	0.965+	1.277***	1.255***
standard error	[0.012]	[0.018]	[0.038]	[0.034]
Pupil is in receipt of Special Educational Needs (SEN) provision: SEN support	0.151***	0.190***	4.135**	4.305***
standard error	[0.022]	[0.034]	[1.117]	[1.001]

	Attainment model	Attainment model - shortlist	SDQ model	SDQ model - shortlist
Pupil is in receipt of Special Educational Needs (SEN) provision: EHCP	0.039***	0.055***	14.146***	9.877***
standard error	[0.015]	[0.022]	[5.818]	[3.333]
Pupil has been in receipt of Free School Meal (FSM) provision within the last 6 years: Yes	0.941	0.761	0.779	0.852
standard error	[0.135]	[0.159]	[0.185]	[0.168]
Pupil classed as a Child in Need (CIN) at time of data collection: Yes	0.596	0.696	2.893	2.413
standard error	[0.179]	[0.398]	[2.249]	[1.667]
Model statistics	-	-	-	-
Num.Obs.	3163	1944	1812	1905
R^2	0.206	0.169	0.333	0.304

Three variables were included in the final list of controls because these had been found in previous research to be important controls for young people's academic attainment, but which in these models were not found to be statistically significant. These were: 1) whether the young person has a health problem (in this analysis, whether they have a health, development, or behavioural issue), whether they receive free school meals, and whether they were a Child in Need. Finally, the SEED disadvantage group was included in these models because of the oversampling of the two most disadvantaged quintiles of children.

Appendix 4: Further attainment models

Table 30: Regression results for childcare hours and grammar, punctuation, and spelling results

	(1) Expected standard in grammar, punctuation, and spelling	(2) Higher standard in grammar, punctuation, and spelling
(Intercept)	2.951**	0.165***
standard error	[0.962]	[0.055]
Childcare hours	-	-
Hours of informal childcare	0.999	1.002
standard error	[0.009]	[0.006]
Hours of formal group childcare	1	1
standard error	[0.007]	[0.006]
Hours of formal individual childcare	1.023	1.001
standard error	[0.015]	[0.010]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**No**	**No**
Model statistics	-	-
Num.Obs.	3325	3325
Pseudo R2 (Cox-Snell)	0.213	0.12
Log.Lik.	-1477.41	-1749.15

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in grammar, punctuation, and spelling at KS2 (model 1) and whether the child had achieved the higher standard in grammar, punctuation, and spelling at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 31: Regression results for childcare hours and maths results

	(1) Expected standard in maths	(2) Higher standard in maths
(Intercept)	4.596***	0.145***
standard error	[1.808]	[0.046]
Childcare hours	-	-
Hours of informal childcare	1.020*	0.993
standard error	[0.010]	[0.006]
Hours of formal group childcare	1.012+	1.004
standard error	[0.007]	[0.006]
Hours of formal individual childcare	1.006	1.01
standard error	[0.011]	[0.009]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**No**	**No**
Model statistics	-	-
Num.Obs.	3325	3325
Pseudo R2 (Cox-Snell)	0.192	0.138
Log.Lik.	-1521.34	-1537.26

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in maths at KS2 (model 1) and whether the child had achieved the higher standard in maths at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 32: Regression results for childcare hours and reading results

	(1) Expected standard in reading	(2) Higher standard in reading
(Intercept)	2.514*	0.265***
standard error	[0.920]	[0.077]
Early childhood education and care predictors	-	-
Hours of informal childcare	1.001	1
standard error	[0.007]	[0.007]
Hours of formal group childcare	1.019*	1.009+
standard error	[0.008]	[0.005]
Hours of formal individual childcare	1.007	0.999
standard error	[0.012]	[0.010]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**No**	**No**
Model statistics	-	-
Num.Obs.	3325	3325
Pseudo R2 (Cox-Snell)	0.189	0.108
Log.Lik.	-1474.71	-1805.69

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in reading at KS2 (model 1) and whether the child had achieved the higher standard in reading at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 33: Regression results for childcare hours and writing results

	(1) Expected standard in writing	(2) Higher standard in writing
(Intercept)	1.796	0.010***
standard error	[0.626]	[0.009]
Early childhood education and care predictors	-	-
Hours of informal childcare	1.004	1.009
standard error	[0.009]	[0.006]
Hours of formal group childcare	1.01	1.005
standard error	[0.007]	[0.007]
Hours of formal individual childcare	1.024+	1.005
standard error	[0.012]	[0.011]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**No**	**No**
Model statistics	-	-
Num.Obs.	3325	1
Pseudo R2 (Cox-Snell)	0.239	1.865**
Log.Lik.	-1459.3	0.984

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in writing at KS2 (model 1) and whether the child had achieved the higher standard in writing at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 34: Regression results for childcare hours and science results

	(1) Expected standard in science
(Intercept)	2.782*
standard error	[1.059]
Early childhood education and care predictors	-
Hours of informal childcare	1.001
standard error	[0.010]
Hours of formal group childcare	1.016
standard error	[0.012]
Hours of formal individual childcare	1.039*
standard error	[0.016]
Controls: Socio-demographic	**Yes**
Controls: Home environment	**No**
Model statistics	-
Num.Obs.	3325
Pseudo R2 (Cox-Snell)	0.27
Log.Lik.	-1137.28

Note: The model presented in the table is a binary logistic regression model displaying odds ratios. Outcome for the model was a binary variable indicating whether the child had achieved the expected standard in science at KS2 (model 1). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 35: Regression results for childcare hours and maths results (scaled score)

	(1) Scaled score in maths
(Intercept)	103.665***
standard error	[0.815]
Early childhood education and care predictors	
Hours of informal childcare	0.011
standard error	[0.018]
Hours of formal group childcare	0.012
standard error	[0.016]
Hours of formal individual childcare	0.035
standard error	[0.025]
Controls: Socio-demographic	**Yes**
Controls: Home environment	**No**
Model statistics	-
Num.Obs.	3200
R2	0.254

Note: The model presented in the table is Ordinary Least Squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcome for the model is a continuous variables indicating the child's scaled score in maths at KS2 (model 1). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 36: Regression results for childcare hours and reading results (scaled score)

	(1) Scaled score in reading
(Intercept)	103.563***
standard error	[0.849]
Early childhood education and care predictors	-
Hours of informal childcare	0.006
standard error	[0.019]
Hours of formal group childcare	0.041*
standard error	[0.018]
Hours of formal individual childcare	0.025
standard error	[0.025]
Controls: Socio-demographic	**Yes**
Controls: Home environment	**No**
Model statistics	-
Num.Obs.	3200
R2	0.231

Note: The model presented in the table is an ordinary least squares regression model. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcome for the model is a continuous variables indicating the child's scaled score in reading at KS2 (model 1). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 37: Regression results for childcare hours and grammar, punctuation, and spelling results (scaled score)

	(1) Scaled score in grammar, punctuation, and spelling (GPS)
(Intercept)	102.717***
standard error	[0.891]
Early childhood education and care predictors	-
Hours of informal childcare	-0.007
standard error	[0.021]
Hours of formal group childcare	0.005
standard error	[0.019]
Hours of formal individual childcare	0.014
standard error	[0.026]
Controls: Socio-demographic	**Yes**
Controls: Home environment	**No**
Model statistics	-
Num.Obs.	3205
R2	0.261

Note: The model presented in the table is an Ordinary Least Squares regression model. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcome for the model is a continuous variables indicating the child's scaled score in in grammar, punctuation, and spelling (GPS) at KS2 (model 1). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 38: Regression results for childcare hours and average score in reading, GPS, and maths

	(1) Average score in reading, GPS, and maths
(Intercept)	103.195***
standard error	[0.750]
Early childhood education and care predictors	-
Hours of informal childcare	0.002
standard error	[0.017]
Hours of formal group childcare	0.019
standard error	[0.015]
Hours of formal individual childcare	0.025
standard error	[0.023]
Controls: Socio-demographic	**Yes**
Controls: Home environment	**No**
Model statistics	-
Num.Obs.	3178
R2	0.279

Note: The model presented in the table is an ordinary least squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcome for the model is a continuous variables indicating the child's scaled score in in Average score in reading, GPS, and maths at KS2 (model 1). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Appendix 5: Exploring the home environment measures

There is significant correlation among the nine home environment measures in the study. Measures which can broadly be understood as capturing potentially problematic home environments are correlated. They are also negatively correlated with measures that indicate a potentially positive home environment (Table 41). This is a limitation of the analysis conducted in chapters 3 and 4, because a high level of correlation between predictors in a regression model can make it difficult to reliably attribute a unique effect on the outcome to each predictor.

Appendix 5 presents findings from a principal component analysis conducted on the nine home environment measures to explore the interrelationships of these concepts. Before entering into the PCA the variables were mean centred and standardised. An overview of the variables distributions and correlations are presented (prior to mean centring and standardisation) on the following page (Table 41).

Prior to conducting the main analysis, a principal component analysis was conducted. This explored the interrelationships among the different measures of home environment including all nine measures for the 3,792 children in the sample at Wave 3. This found that the first three principal components explain 56% of the variance in these items. This is a reasonable amount but not very high, which suggests the data will not naturally reduce into very few dimensions. In practice this means that these components are only able to partially explain the original data; the individual measures of the home environment. The first component explains 32% of the variance, component 2 13% and component 3 11%. After that the remaining components explain steadily less (between 9% and 5%).

Component 1 appears to be an overall measure of the home environment's quality, although more strongly related to the negative elements of the home environment. Children with low scores on it would have poorer home environments, and people with higher scores a better home environment. It was most strongly (and negatively) correlated with the Mother Object Relations Scales (MORS) invasiveness score, the authoritarian and permissive parenting scores, as well as how chaotic the household was (the Confusion, Hubbub, and Order Scale - CHAOS). It was also more weakly and positively correlated with the authoritative parenting score, the MORS warmth scale, and Home Learning Environment (HLE) index.

That component 1 is negatively correlated both with the parental authoritarianism and permissiveness scores sounds counterintuitive, however, the underlying items for these scores are linked. The authoritarian score measures the frequency of behaviours such as shouting at, scolding or criticising the child, becoming angry with them, or using physical punishments. While the permissiveness score measures parental behaviours such as difficulties in disciplining their child or giving into their child.

By contrast, component 2 was most strongly related to three positive aspects of the home environment: the warmth of the parental relationship, the quality of the home learning environment, and the authoritative parenting score. The authoritative parenting score measures items such as encouraging the child to talk about their feelings, explaining the consequences of their actions, or giving praise when the child behaves well. This component then provides a summary of these positive measures of the home environment, while not being strongly related to the negative measures (such as the MORS invasiveness scale).

Finally, component 3 was strongly related to one item, the parental mental distress score, and was comparatively weakly related to the other items. This would suggest that the parental mental distress score captures something more unique about the home environment, which is distinctive to the other measures. Another item that did not strongly fit with any of these three components in particular was parental limit setting.

Overall, the components seem to identify two important underlying dimensions of these nine home environment measures, suggesting that there are important commonalities between the concepts they measure, which is also reflected in the correlations between items. However, it also suggests that these dimensions are not that clearly defined, because the correlations between individual items and the components are not very strong. To conclude, when considering each of these home environment measures in isolation, it is important to be aware of these overlaps and that confounding factors could be present.

Table 39: Summary of components

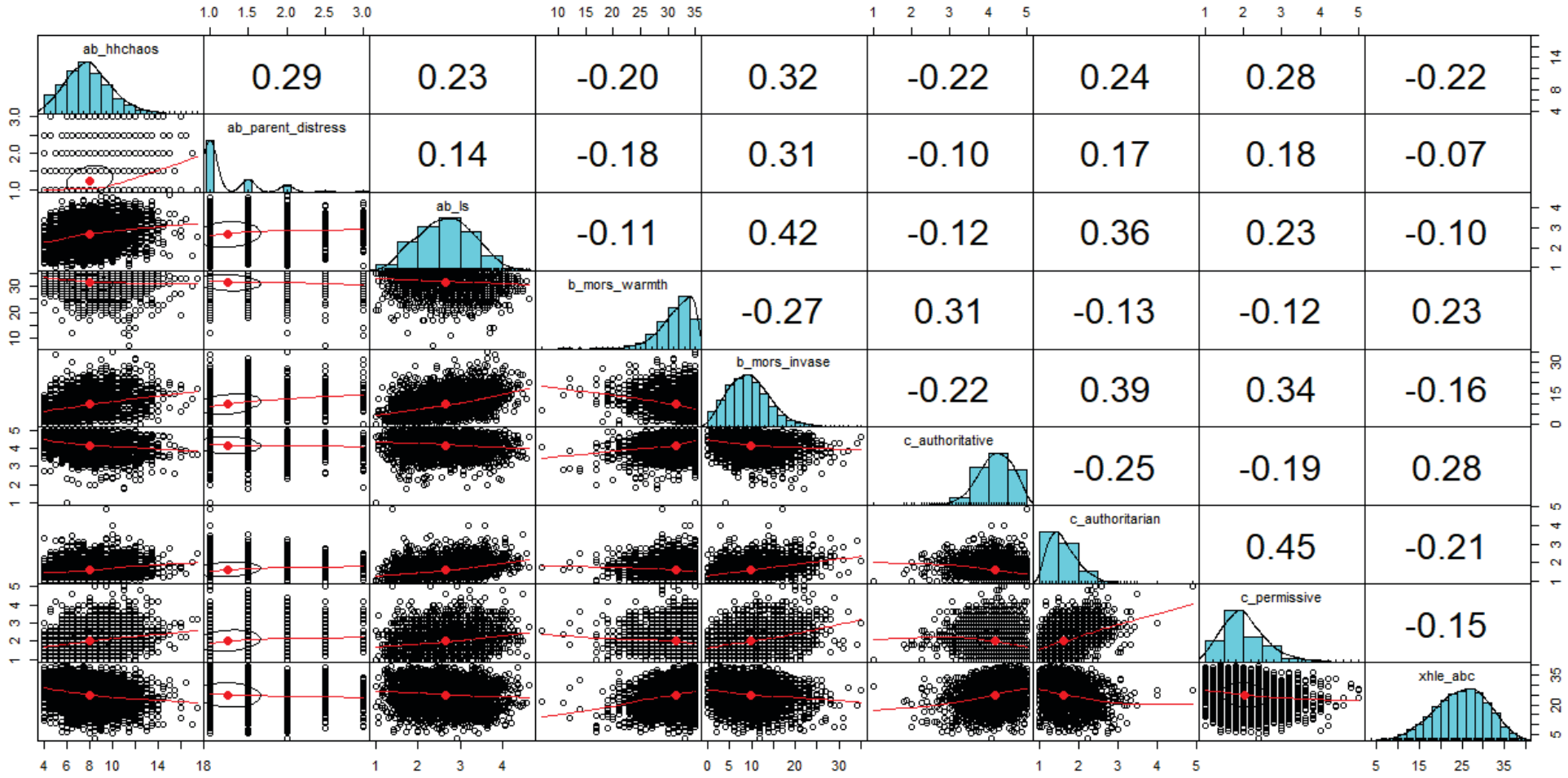
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Standard deviation	1.70	1.09	1.00	0.92	0.88	0.83	0.81	0.72	0.70
Proportion of Variance	0.32	0.13	0.11	0.09	0.09	0.08	0.07	0.06	0.05
Cumulative Proportion	0.32	0.45	0.56	0.66	0.74	0.82	0.89	0.95	1.00

Table 40: Component loadings

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Household chaos score	-0.35	0.04	0.28	-0.33	-0.38	0.62	0.32	-0.08	0.21
Parental distress	-0.27	-0.08	0.71	-0.26	0.07	-0.3	-0.45	-0.17	-0.15
Parental limit setting	-0.33	-0.37	-0.11	0.56	-0.38	0.1	-0.1	-0.31	-0.41

MORS warmth scale	0.28	-0.5	-0.24	-0.41	-0.29	0.2	-0.49	0.3	-0.04
MORS invasiveness scale	-0.42	-0.2	0.14	0.3	0.02	-0.11	0.04	0.78	0.24
Authoritative parenting score	0.3	-0.47	0.21	-0.07	-0.32	-0.47	0.53	-0.1	0.16
Authoritarian parenting score	-0.4	-0.25	-0.37	-0.1	0.17	-0.17	-0.17	-0.39	0.63
Permissive parenting score	-0.36	-0.25	-0.26	-0.43	0.38	-0.08	0.36	0.07	-0.52
HLE index	0.26	-0.48	0.29	0.24	0.58	0.45	0.06	-0.11	0.1

Table 41: Home environment measures overview[6]



[6] This table shows an overview of each of the home environment variables' distributions in the cells that appear diagonally from top-left to bottom-right. In the top right-hand half of the table are the correlations of each pair of variables, and in the bottom-left are plots of the distributions of each pair of variables by each other.

Appendix 6: Complete results chapter regression tables

This appendix includes the complete version of the results chapter regression tables. For clarity of reporting the tables in the main body of the report excluded the coefficient standard errors and model intercepts shown below.

Table 42 Complete regression results table for overall attainment and childcare hours

	(1) Expected standard	(2) Expected standard	(3) Higher standard	(4) Higher standard
(Intercept)	1.065	0.433	0.000***	0.000***
standard error	[0.351]	[0.380]	[0.000]	[0.000]
Childcare hours	-	-	-	-
Hours of informal childcare	1.004	1.004	1.007	1.005
standard error	[0.007]	[0.007]	[0.010]	[0.010]
Hours of formal group childcare	1.017*	1.015*	1.018*	1.020*
standard error	[0.007]	[0.007]	[0.008]	[0.008]
Hours of formal individual childcare	1.004	1.003	1	1.003
standard error	[0.009]	[0.009]	[0.013]	[0.014]
Controls: Socio-demographic	**Yes**	**Yes**	**Yes**	**Yes**
Controls: Home environment	**No**	**Yes**	**No**	**Yes**
Model statistics	-	-	-	-
Num.Obs.	3325	3232	3325	3232
Pseudo R2 (Cox-Snell)	0.188	0.195	0.132	0.144
Log.Lik.	-1785.625	-1712.1	-772.129	-750.919

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in reading, writing, and maths at KS2 (models 1 and 2) and whether the child had achieved the higher expected standard in reading, writing, and maths at KS2 (models 3 and 4). The

standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 43 Complete regression results table for overall attainment and childcare hours, stratified by disadvantage

	(1) Full sample	(2) Most deprived	(3) 2nd quintile	(4) Least deprived
(Intercept)	1.244	0.493+	2.722*	0.965
standard error	[0.398]	[0.189]	[1.260]	[0.779]
Childcare hours:	-	-	-	-
Hours of informal childcare	1.007	1.021	1.007	1
standard error	[0.007]	[0.020]	[0.011]	[0.009]
Hours of formal group childcare	1.017*	1.045*	1.001	1.020+
standard error	[0.007]	[0.021]	[0.011]	[0.010]
Hours of formal individual childcare	1.004	1.046	0.995	1.005
standard error	[0.010]	[0.053]	[0.014]	[0.017]
Controls: Socio-demographic	Yes	Yes	Yes	Yes
Controls: Home environment	No	No	No	No
Num.Obs.	3325	834	1201	1290
Pseudo R2 (Cox-Snell)	0.185	0.194	0.166	0.162

Note: Outcomes for model was Achieved expected standard in reading, writing, and maths at KS2. Model in the table is a binary logistic regression models and the table is displaying odds ratios. The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 44 Complete regression results table for overall attainment and age started formal ECEC

	(1) Expected standard	(2) Higher stand-ard
(Intercept)	0.34	0.000***
standard error	[0.320]	[0.000]
Age started nursery:	-	-
Before age 1	1.768	1.613
standard error	[0.633]	[0.966]
Between 1 and 2	2.061+	1.689
standard error	[0.762]	[1.090]
Between 2 and 3	1.727	0.982
standard error	[0.627]	[0.640]
Between 3 and 4	1.492	1.384
standard error	[0.530]	[0.895]
After the age of 4	1.465	1.37
standard error	[0.522]	[0.848]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**Yes**	**Yes**
Model statistics	-	-
Num.Obs.	3232	3232
Pseudo R2 (Cox-Snell)	0.195	0.143
Log.Lik.	-107.32	-114.33

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in reading, writing, and maths at KS2 (model 1) and whether the child had achieved the higher expected standard in reading, writing, and maths at KS2 (model 2). 'Age started' displays the coefficients for the categorical variable showing the age at which the child first received more than 10 hours a week for formal ECEC (on average), with children who never received more than 10 hours a week of formal ECEC constituting the reference category. The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 45 Complete regression results table for overall attainment and ECEC setting quality

	(1) Expected standard	(2) Higher stand-ard
(Intercept)	0.5	0.000***
standard error	[0.438]	[0.000]
Setting quality:	-	-
Good	1.063	0.986
standard error	[0.238]	[0.301]
Excellent	0.848	0.908
standard error	[0.241]	[0.373]
No setting quality measure	0.948	1.562
standard error	[0.167]	[0.448]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**Yes**	**Yes**
Model statistics	-	-
Num.Obs.	3232	3232
Pseudo R2 (Cox-Snell)	0.194	0.146
Log.Lik.	-1714.1	-748.51

Note: Models presented in the table are binary logistic regression models. The coefficients in the table are displayed as odds ratios. Outcomes for models were binary variables indicating whether the child had achieved the expected standard in reading, writing, and maths at KS2 (model 1) and whether the child had achieved the higher expected standard in reading, writing, and maths at KS2 (model 2). 'Setting quality' displays the SSTEW measure for the setting the child attended recorded as a categorical variable, with the rating minimal or inadequate constituting the reference category. The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 46 Complete regression results table for socio-emotional outcomes and childcare hours

	(1) Total difficulties score (W7)	(2) Total difficulties score (W7)	(3) Total difficulties score (W7)	(4) Prosocial score (W7)	(5) Prosocial score (W7)	(6) Prosocial score (W7)
(Intercept)	2.810+	-2.334	10.779***	9.344***	7.469***	7.932***
standard error	[1.422]	[3.301]	[1.311]	[0.554]	[0.927]	[0.472]
Childcare hours:	-	-	-	-	-	-
Hours of informal childcare	0.007	0.005	0.014	-0.006	-0.006	-0.007
standard error	[0.015]	[0.015]	[0.018]	[0.005]	[0.005]	[0.005]
Hours of formal group childcare	0.035*	0.032*	0.011	-0.019*	-0.017*	-0.014+
standard error	[0.014]	[0.015]	[0.014]	[0.007]	[0.007]	[0.007]
Hours of formal individual childcare	0.046*	0.046*	0.017	-0.012	-0.015+	-0.007
standard error	[0.022]	[0.022]	[0.025]	[0.009]	[0.009]	[0.009]
Controls: Socio-demographic	**Yes**	**Yes**	**Yes** Excl. SDQ Wv2/3	**Yes**	**Yes**	**Yes** Excl. SDQ Wv2/3
Controls: Home environment	**No**	**Yes**	**No**	**No**	**Yes**	**No**
Model statistics	-	-	-	-	-	-

Num.Obs.	1905	1862	1907	1913	1869	1915
R2	0.412	0.425	0.27	0.213	0.228	0.154

Note: Models presented in the table are Ordinary Least Squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcomes for models were continuous variables indicating total difficulties score at KS2 (models 1 and 2) and prosocial score at KS2 (models 3 and 4). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, *

Table 47 Complete regression results table for the SDQ subscales and childcare hours

	(1) Conduct problems	(2) Emotional problems	(3) Hyper-activity problems	(4) Peer problems
(Intercept)	0.398	0.436	1.133*	0.886+
standard error	[0.432]	[0.475]	[0.471]	[0.457]
Childcare hours:	-	-	-	-
Hours of informal childcare	0.006	0.001	0.001	-0.001
standard error	[0.005]	[0.007]	[0.008]	[0.005]
Hours of formal group childcare	0.013**	0.009	0.014+	-0.001
standard error	[0.005]	[0.007]	[0.008]	[0.005]
Hours of formal individual childcare	0.017*	0.017	0.017+	-0.004
standard error	[0.007]	[0.011]	[0.010]	[0.007]
Controls: Socio-demographic	**Yes**	**Yes**	**Yes**	**Yes**
Controls: Home environment	**No**	**No**	**No**	**No**
Model statistics	-	-	-	-
Num.Obs.	1912	1914	1910	1911
R^2	0.237	0.227	0.307	0.271

Note: Models presented in the table are Ordinary Least Squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcomes for models were continuous variables indicating SDQ subscales. The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, * denotes p ≤ 0.05, ** denotes p ≤ 0.01, and *** denotes p ≤ 0.001.

Table 48 Complete regression results table for socio-emotional outcomes and age started receiving childcare

	(1) Total difficulties score (W7)	(2) Prosocial score (W7)
(Intercept)	3.531	6.192***
standard error	[3.747]	[0.900]
Age started nursery:	-	-
Before age 1	1.577	-0.193
standard error	[1.191]	[0.378]
Between 1 and 2	1.651	-0.123
standard error	[1.245]	[0.405]
Between 2 and 3	1.478	0.091
standard error	[1.335]	[0.368]
Between 3 and 4	1.676	0.075
standard error	[1.230]	[0.390]
After the age of 4	1.467	0.05
standard error	[1.188]	[0.402]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**Yes**	**Yes**
Model statistics	-	-
Num.Obs.	1862	1869
R^2	0.349	0.205

Note: Models presented in the table are Ordinary Least Squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcomes for models were continuous variables indicating total difficulties score at KS2 (model 1) and prosocial score at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, *

Table 49 Complete regression results table for socio-emotional outcomes and ECEC setting quality

	(1) Total difficulties score (W7)	(2) Prosocial score (W7)
(Intercept)	4.916	6.244***
standard error	[3.550]	[0.901]
Setting quality:	-	-
Good	-0.045	-0.223
standard error	[0.528]	[0.182]
Excellent	1.293	0.175
standard error	[0.779]	[0.191]
No setting quality measure	0.155	-0.043
standard error	[0.461]	[0.180]
Controls: Socio-demographic	**Yes**	**Yes**
Controls: Home environment	**Yes**	**Yes**
statistics	-	-
Num.Obs.	1862	1869
R^2	0.349	0.205

Note: Models presented in the table are Ordinary Least Squares regression models. The coefficients represent the expected change in the outcome variable for each one-unit increase in that predictor, holding all other variables constant. Outcomes for models were continuous variables indicating total difficulties score at KS2 (model 1) and prosocial score at KS2 (model 2). The standard error of each coefficient is shown below it in brackets. Levels of statistical significance are denoted as: + p <= 0.1, *

Table 50 Complete regression results table for achieving the expected standard in Reading, Writing, and Maths on children's early years home environment

	(1) Achieved expected standard	(2) Achieved expected standard	(3) Achieved expected standard
(Intercept)	0.241*	0.285	0.329
standard error	[0.153]	[0.333]	[0.311]

	(1) Achieved expected standard	(2) Achieved expected standard	(3) Achieved expected standard
How far parents put limits on their children's behaviour (average W1 and W2)	1.330***	1.444***	1.200*
standard error	[0.091]	[0.125]	[0.096]
CHAOS scale - average for W1, W2	0.925***	-	1.003
standard error	[0.020]	-	[0.026]
Parental mental distress (Waves 1 and 2)	1.027+	-	0.986
standard error	[0.014]	-	[0.016]
MORS invasiveness subscale (W2)	0.992	0.994	1.016
standard error	[0.013]	[0.015]	[0.015]
MORS warmth subscale (W2)	0.997	0.98	0.979
standard error	[0.014]	[0.024]	[0.017]
Authoritarian parenting score (W3)	0.949	-	0.88
standard error	[0.109]	-	[0.129]
Authoritative parenting score (W3)	1.225+	-	1.309*
standard error	[0.124]	-	[0.156]
Permissive parenting score (W3)	0.834*	-	0.984
standard error	[0.063]	-	[0.087]
HLE index (average over W1, W2 and W3)	1.033***	1.039**	1.022*
standard error	[0.009]	[0.012]	[0.009]
CHAOS scale - average for W1, W2 and W7	-	0.912*	-
standard error	-	[0.031]	-
Parental mental distress (W1, W2 and W7)	-	1.054**	-
standard error	-	[0.020]	-
Authoritarian parenting score (average W3 and W7)	-	1.023	-
standard error	-	[0.233]	-

	(1) Achieved expected standard	(2) Achieved expected standard	(3) Achieved expected standard
Authoritative parenting score (average W3 and W7)	-	1.207	-
standard error	-	[0.188]	-
Permissive parenting score (average W3 and W7)	-	0.648**	-
standard error		[0.093]	
Controls: Socio-demographic	**No**	**No**	**Yes**
Controls: ECEC measures	**No**	**No**	**Yes: childcare hours, quality and age started**
Num.Obs.	3246	1922	3232
Pseudo R^2 (Cox-Snell)	0.029	0.048	0.196

Table note: The home environment variables are strongly correlated. Children with concerning scores on the Home Learning Environment index also tend to have concerning scores on the CHAOS scale, for example. Correlated predictors in a regression model means that the individual coefficients should not be interpreted, as the interpretation risks being misleading. In more technical terms, the correlated predictors cause inflated standard errors for the estimated coefficients, making it difficult to determine the individual effect of correlated predictors. In these models, the variance inflation factors (VIF) for each of the coefficients are all above 5 and many are above 20. The model coefficients for limit setting behaviour, the CHAOS score etc will therefore not be interpreted, but as they were made available in the SEED KS1 report they are included for comparison purposes.

Table 51 Complete regression results table for achieving greater than the expected standard in Reading, Writing, and Maths on children's early years home environment

	(1) Achieved greater than expected standard	(2) Achieved greater than expected standard	(3) Achieved greater than expected standard
Intercept	0.079+	0.287	0.000***
standard error	[0.102]	[0.570]	[0.000]
How far parents put limits on their children's behaviour (average W1 and W2)	1.146	1.04	0.914
standard error	[0.137]	[0.133]	[0.123]
CHAOS scale - average for W1, W2	0.905*	-	0.989
standard error	[0.038]	-	[0.043]
Parental mental distress (Waves 1 and 2)	1.042	-	1.026
standard error	[0.028]	-	[0.034]
MORS invasiveness subscale (W2)	0.97	1.006	0.986
standard error	[0.017]	[0.024]	[0.021]
MORS warmth subscale (W2)	0.937**	0.929*	0.933**
standard error	[0.018]	[0.026]	[0.022]
Authoritarian parenting score (W3)	0.807	-	0.927
standard error	[0.172]	-	[0.257]
Authoritative parenting score (W3)	1.153	-	1.181
standard error	[0.180]	-	[0.215]
Permissive parenting score (W3)	1.087	-	1.340*
standard error	[0.109]	-	[0.174]
HLE index (average over W1, W2 and W3)	1.051***	1.053***	1.042**
standard error	[0.012]	[0.014]	[0.013]
CHAOS scale - average for W1, W2 and W7	-	0.879+	-
standard error	-	[0.059]	-

	(1) Achieved greater than expected standard	(2) Achieved greater than expected standard	(3) Achieved greater than expected standard
Parental mental distress (W1, W2 and W7)	-	1.063+	-
standard error	-	[0.037]	-
Authoritarian parenting score (average W3 and W7)	-	0.511*	-
standard error	-	[0.169]	-
Authoritative parenting score (average W3 and W7)	-	1.082	-
standard error	-	[0.283]	-
Permissive parenting score (average W3 and W7)	-	0.94	-
standard error	-	[0.173]	-
Controls: Socio-demographic	**No**	**No**	**Yes**
Controls: ECEC measures	**No**	**No**	**Yes: childcare hours, quality and age started**
Num.Obs.	3246	1922	3232
Pseudo R^2 (Cox-Snell)	0.029	0.044	0.153

Table note: The home environment variables are strongly correlated. Children with concerning scores on the Home Learning Environment index also tend to have concerning scores on the CHAOS scale, for example. Correlated predictors in a regression model means that the individual coefficients should not be interpreted, as the interpretation risks being misleading. In more technical terms, the correlated predictors cause inflated standard errors for the estimated coefficients, making it difficult to determine the individual effect of correlated predictors. In these models, the variance inflation factors (VIF) for each of the coefficients are all above 5 and many are above 20. The model coefficients for limit setting behaviour, the CHAOS score etc will therefore not be interpreted, but as they were made available in the SEED KS1 report they are included for comparison purposes.

Table 52 Complete regression results table for the home environment and total difficulties score

	(1) Total difficulties score (W7)	(2) Total difficulties score (W7)	(3) Total difficulties score (W7)	(4) Total difficulties score (W7)
Intercept	11.410**	11.183**	-4.848	3.279
standard error	[3.790]	[3.490]	[3.720]	[3.841]
How far parents put limits on their children's behaviour (average W1 and W2)	0.254	-0.209	0.504	0.827*
standard error	[0.340]	[0.308]	[0.333]	[0.329]
CHAOS scale - average for W1, W2	0.433***	-	0.128	0.219*
standard error	[0.083]	-	[0.079]	[0.085]
Parental mental distress (Waves 1 and 2)	-0.259***	-	0.011	-0.092
standard error	[0.065]	-	[0.053]	[0.056]
MORS invasiveness subscale (W2)	0.208***	0.089	0.013	0.177***
standard error	[0.056]	[0.057]	[0.042]	[0.045]
MORS warmth subscale (W2)	-0.085	-0.051	0.064	-0.043
standard error	[0.072]	[0.064]	[0.057]	[0.058]
Authoritarian parenting score (W3)	0.913	-	0.384	0.956+
standard error	[0.550]	-	[0.443]	[0.519]
Authoritative parenting score (W3)	-0.134	-	-0.072	-0.086

	(1) Total difficulties score (W7)	(2) Total difficulties score (W7)	(3) Total difficulties score (W7)	(4) Total difficulties score (W7)
standard error	[0.352]	-	[0.354]	[0.363]
Permissive parenting score (W3)	0.511	-	-0.159	0.167
standard error	[0.339]	-	[0.329]	[0.354]
HLE index (average over W1, W2 and W3)	-0.008	0.019	0.031	0.046
standard error	[0.033]	[0.027]	[0.027]	[0.027]
CHAOS scale - average for W1, W2 and W7	-	0.694***	-	-
standard error	-	[0.099]	-	-
Parental mental distress (W1, W2 and W7)	-	-0.478***	-	-
standard error	-	[0.063]	-	-
Authoritarian parenting score (average W3 and W7)	-	2.390**	-	-
standard error	-	[0.692]	-	-
Authoritative parenting score (average W3 and W7)	-	-0.109	-	-
standard error	-	[0.350]	-	-
Permissive parenting score (average W3 and W7)	-	1.495***	-	-
standard error	-	[0.365]	-	-

	(1) Total difficulties score (W7)	(2) Total difficulties score (W7)	(3) Total difficulties score (W7)	(4) Total difficulties score (W7)
Controls: Socio-demographic	**No**	**No**	**Yes**	**Yes** Excl. SDQ Wv2/3
Controls: ECEC	**No**	**No**	**Yes**	**Yes**
Model statistics	-	-	-	-
Num.Obs.	2049	2074	1862	1862
R^2	0.143	0.256	0.429	0.351

Table note: The home environment variables are strongly correlated. Children with concerning scores on the Home Learning Environment index also tend to have concerning scores on the CHAOS scale, for example. Correlated predictors in a regression model means that the individual coefficients should not be interpreted, as the interpretation risks being misleading. In more technical terms, the correlated predictors cause inflated standard errors for the estimated coefficients, making it difficult to determine the individual effect of correlated predictors. In these models, the variance inflation factors (VIF) for each of the coefficients are all above 5 and many are above 20. The model coefficients for limit setting behaviour, the CHAOS score etc will therefore not be interpreted, but as they were made available in the SEED KS1 report they are included for comparison purposes.

Table 53 Complete regression results table for the home environment and prosocial score

	(1) Prosocial score (W7)	(2) Prosocial score (W7)	(3) Prosocial score (W7)	(4) Prosocial score (W7)
Intercept	5.312***	3.899**	7.120***	6.202***
standard error	[1.061]	[1.280]	[1.007]	[0.914]
How far parents put limits on their children's behaviour (average W1 and W2)	-0.037	0.041	-0.129	-0.162

	(1) Prosocial score (W7)	(2) Prosocial score (W7)	(3) Prosocial score (W7)	(4) Prosocial score (W7)
standard error	[0.090]	[0.082]	[0.097]	[0.094]
CHAOS scale - average for W1, W2	-0.038	-	0.003	-0.008
standard error	[0.026]	-	[0.027]	[0.026]
Parental mental distress (Waves 1 and 2)	0.037*	-	-0.008	0.008
standard error	[0.017]	-	[0.015]	[0.016]
MORS invasiveness subscale (W2)	-0.052***	-0.037**	-0.021	-0.045**
standard error	[0.014]	[0.012]	[0.014]	[0.015]
MORS warmth subscale (W2)	0.059**	0.037*	0.035*	0.051**
standard error	[0.019]	[0.018]	[0.015]	[0.015]
Authoritarian parenting score (W3)	0.09	-	0.169	0.072
standard error	[0.186]	-	[0.172]	[0.178]
Authoritative parenting score (W3)	0.204+	-	0.194+	0.215+
standard error	[0.111]	-	[0.100]	[0.107]
Permissive parenting score (W3)	-0.142	-	-0.008	-0.061
standard error	[0.107]	-	[0.096]	[0.096]
HLE index (average over W1, W2 and W3)	0.020*	0.011	0.008	0.005
standard error	[0.008]	[0.008]	[0.008]	[0.008]

	(1) Prosocial score (W7)	(2) Prosocial score (W7)	(3) Prosocial score (W7)	(4) Prosocial score (W7)
CHAOS scale - average for W1, W2 and W7	-	-0.081*	-	-
standard error	-	[0.031]	-	-
Parental mental distress (W1, W2 and W7)	-	0.055**	-	-
standard error	-	[0.020]	-	-
Authoritarian parenting score (average W3 and W7)	-	-0.109	-	-
standard error	-	[0.219]	-	-
Authoritative parenting score (average W3 and W7)	-	0.731***	-	-
standard error	-	[0.136]	-	-
Permissive parenting score (average W3 and W7)	-	-0.157	-	-
standard error	-	[0.122]	-	-
Controls: Demographics	**No**	**No**	**Yes**	**Yes** Excl. SDQ W2/3
Controls: ECEC Measures	**No**	**No**	**Yes**	**Yes**
Model statistics	-	-	-	-
Num.Obs.	2056	2081	1869	1869
R^2	0.097	0.14	0.234	0.209

Table note: The home environment variables are strongly correlated. Children with concerning scores on the Home Learning Environment index also tend to have

concerning scores on the CHAOS scale, for example. Correlated predictors in a regression model means that the individual coefficients should not be interpreted, as the interpretation risks being misleading. In more technical terms, the correlated predictors cause inflated standard errors for the estimated coefficients, making it difficult to determine the individual effect of correlated predictors. In these models, the variance inflation factors (VIF) for each of the coefficients are all above 5 and many are above 20. The model coefficients for limit setting behaviour, the CHAOS score etc will therefore not be interpreted, but as they were made available in the SEED KS1 report they are included for comparison purposes.

Department for Education

Reference: RR1527

ISBN: 978-1-83870-651-7

For any enquiries regarding this publication, contact www.education.gov.uk/contactus.

This document is available for download at www.gov.uk/government/publications.