Department for Education

IELS thematic report: Young children's development and deprivation in England

Research report

June 2021

Rachel Classick, Chris Hope, Caroline Sharp: National Foundation for Educational Research

Government
Social Research

Contents

List of figures

List of tables

Acknowledgements

The authors would like to thank all the children, parents, teachers and schools that took part in IELS for their time and invaluable contribution to this study. We would also like to thank our National Advisory Committee[1] for their help and guidance throughout this work; and colleagues in the international evidence and statistics and early years research and policy teams at the Department for Education, including Frances Forsyth, Ali Pareas and Jonathon Blackburn.

We would like to thank the OECD and the Australian Council for Educational Research (ACER), the IEA Data processing and Research Center (IEA DPC) and cApStAn Linguistic Quality Control for developing and coordinating the study.

We would also like to thank our NFER colleagues including our statistical team, especially Simon Rutt, Charalambos Kollias and Karim Badr; our colleagues in research operations and IT support; and our project administrator, Neelam Basi, for all their support and assistance in data collection, analysis and reporting.

Finally, we would like to thank Professor Iram Siraj and Professor Kathy Sylva for their expert advice, support and comments on this report.

[1] The role of the National Advisory Committee was to provide independent advice, guidance and constructive feedback to DfE and NFER on the development and implementation of the study.

IELS outputs overview

The main reports produced for IELS are listed below.

Reports published by OECD

- Early learning and child well-being: A study of 5-year-olds in England, Estonia and the United States (OECD, 2020a). This report looks at the findings as a whole and compares and contrasts the findings across the 3 countries.
- Early learning and child well-being in England (OECD, 2020b). This report focuses on the findings in England.
- Early learning and child well-being in Estonia (OECD, 2020c). This report focuses on the findings for IELS in Estonia
- Early learning and child well-being in the United States (OECD, 2020d). This report focuses on the findings for IELS in the United States
- Early International Early Learning and Child Wellbeing Study (IELS) Technical Report (OCED 2021). This report provides technical information about the study and data analysis.

Reports published by NFER

- IELS summary report (Kettlewell and others, 2020a), which summarises findings from the IELS national report for England.
- IELS participant report (Kettlewell and others, 2021), which summarises findings for parents and staff in participating schools.

Reports published by Department for Education

- IELS national report for England (Kettlewell and others, 2020b), which builds on the OECD country report for England by further contextualising the findings for England by linking the IELS data with the national pupil database (NPD) and reporting on national questions and an additional measure of physical development.
- IELS thematic report on disadvantage (this report), which focuses on a more in-depth analysis of the IELS findings on the relationship between children's outcomes and socio-economic disadvantage.
- IELS thematic report on physical development (Lucas and others, 2021), which focuses on a more in-depth investigation of the findings on children's physical development.

Executive Summary

Introduction

The International Early Learning and Child Well-being Study (IELS) is a new study conducted by the OECD. It seeks to understand children's level of development at age 5, and how this is influenced by their early education experiences, the home learning environment (HLE), and individual and family characteristics.

Few studies are able to account for the complex interplay between individual and area when investigating socio-economic status (Crenna-Jennings, 2018). IELS provides an opportunity to explore which of 5 deprivation measures best capture the relationship of deprivation with young children's development and which are the key risk and protective factors in emergent literacy (which comprised oral language rather than reading or writing), emergent numeracy, mental flexibility (children's ability to shift their thinking according to the circumstances) and emotion identification (an important aspect of empathy).

This analysis had 2 aims:

1. To identify how different measures of socio-economic deprivation and school deprivation are related to early learning outcomes at age 5.
2. To identify the protective and risk factors in relation to early learning outcomes at age 5 for all children and explore if these are different for those children from more deprived backgrounds.

The sample comprised 2,577 children aged between 4 years 11 months and 6 years 0 months, from 191 schools. The analysis used multi-level multivariate analysis to identify the relationship between a specific variable and learning outcome, taking account of the influence of other variables including child and family characteristics (such as age, gender, having an identified special educational need (SEN), having English as an additional language (EAL) and deprivation); aspects of the HLE; and children's attention and persistence.

How do different measures of deprivation interact to predict early learning outcomes at age 5?

Parental socio-economic status and school-level measures of deprivation are more strongly related to children's development than child-level measures

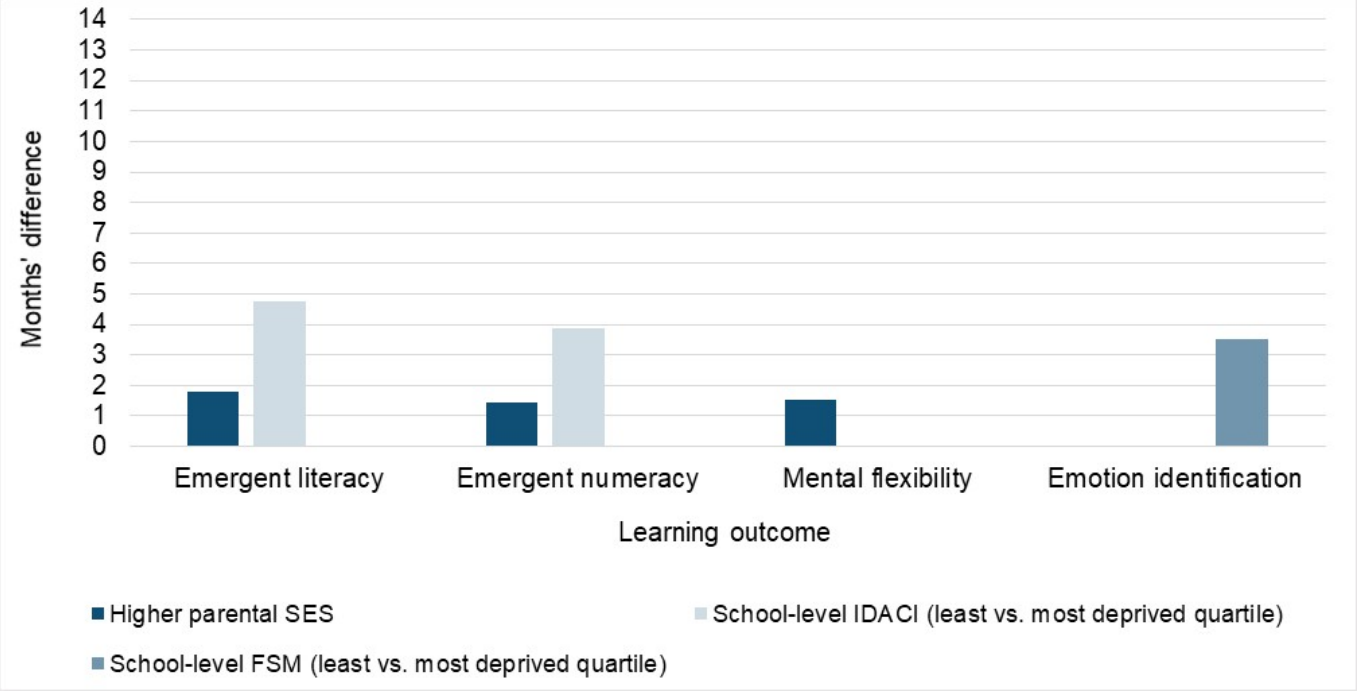
The 5 measures of deprivation used to investigate the relationship between deprivation and learning outcomes were:

- child-level free school meals (FSM)
- school-level FSM
- child-level IDACI
- school-level IDACI
- parental socio-economic status (SES).

Parental SES (comprising parents' occupation, level of education and household income) was related to 3 outcomes (emergent literacy, emergent numeracy and mental flexibility)[2]; the proportion of income-deprived families in the area of the child's school (school-level IDACI) was related to 2 outcomes (emergent literacy and emergent numeracy); and the proportion of pupils eligible for FSM in a child's school was related to one (emotion identification).

Figure 1 shows the differences between the groups of most and least deprived children in each measure, translated into difference in months.

Figure 1: Months' difference in 5-year-olds' learning outcomes by deprivation measure



Source: IELS assessment of 2,577 children, age 5

[2] As SES is a continuous variable, the effect size calculated represents the effect of a change in the independent variable (that is, SES score) equivalent to its adjusted standard deviation on the outcome of interest.

Of the 5 deprivation measures analysed, the parental SES measure is most consistently related to children's learning outcomes. This research did not find a significant association between either child-level FSM or child-level IDACI in any of the 4 learning outcomes after taking account of the influence of other variables. The binary nature of FSM, in comparison to the continuous measure of parental SES, makes it a less sensitive measure of deprivation (Taylor, 2018). As suggested by previous research (Sutherland and others, 2015; Ilie and others, 2017; Taylor, 2018), the combination of parental education, employment and income for the parental SES index makes it a more powerful predictor of the relationship between deprivation and children's outcomes.

At the school level, IDACI had greater predictive power than school-level FSM for emergent literacy and emergent numeracy. However, school-level FSM was related to emotion identification, which is likely to be due to the characteristics of a school's pupil intake which can influence children's social outcomes (Sylva and others, 2008).

Although parental SES was related to 3 learning outcomes, the relationship between school-level IDACI and emergent literacy and emergent numeracy development was stronger (in terms of effect size and the associated months' difference). This suggests that the deprivation of the school influences children's cognitive development, possibly through the quality of teaching and learning children receive and interaction with their peer group at school.

What are the protective and risk factors associated with early learning outcomes at age 5?

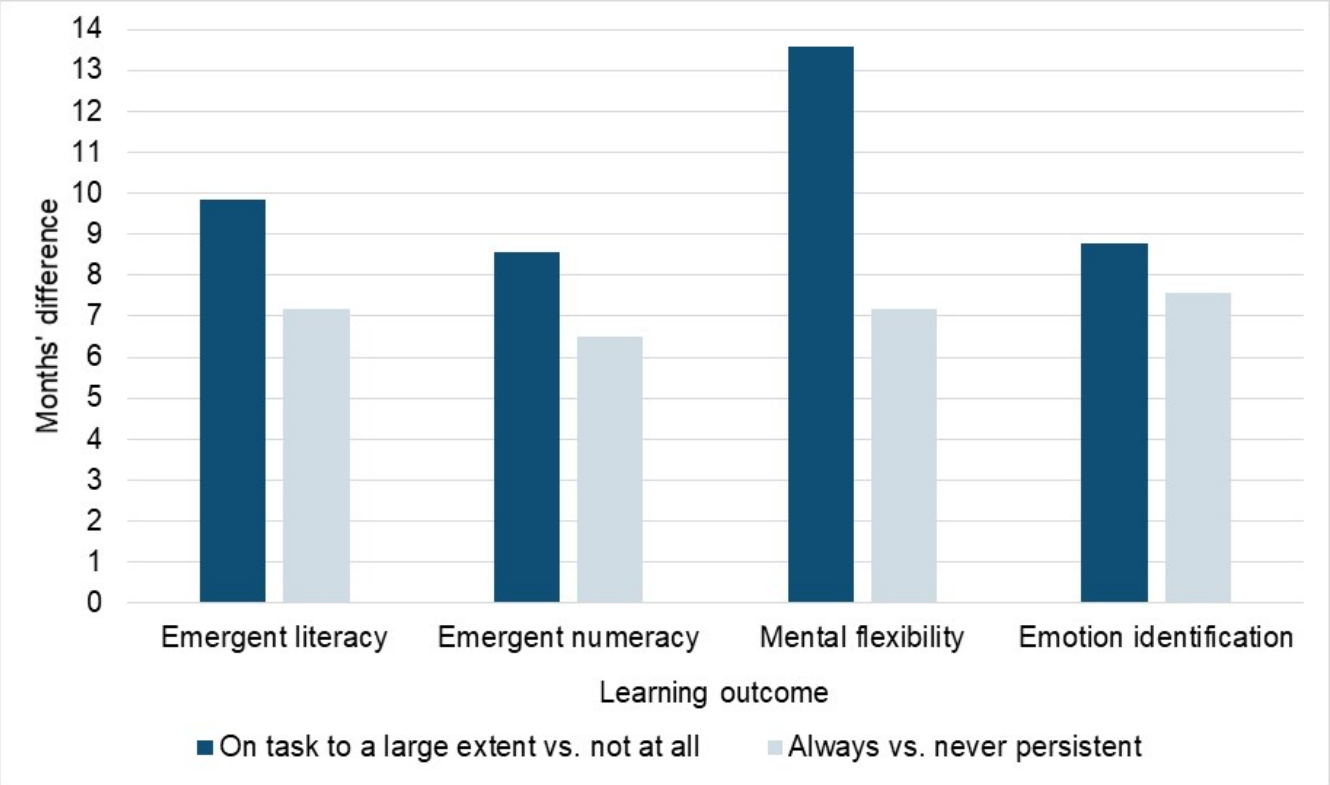
Characteristics that are positively related to children's development can be thought of as protective factors and those that are negatively related can be considered to be risk factors. Although deprivation was a risk factor for all 4 learning outcomes, it did not have the largest effect. This indicates that children's learning outcomes are influenced by a wide range of factors, of which deprivation is just one.

Being on task/not distracted and being more persistent were strongly related to greater development in emergent literacy, emergent numeracy, mental flexibility and emotion identification. Having an identified SEN was a significant risk factor for all 4 learning outcomes.

Attention and persistence are strongly related to children's development

Children who were on task to a large extent had greater development in all 4 learning outcomes compared with children who were not on task at all, as did children who were always persistent compared with those who were never persistent.

Figure 2: Months' difference in 5-year-olds' learning outcomes by attention and persistence level



Source: IELS assessment of 2,577 children, age 5

IELS study administrators rated 2 aspects of children's behaviour during the direct assessments: 'Was the child easily distracted?' and 'Did the child stay on task?'. Being on task to a large extent (compared to not being on task at all) was more strongly related to children's learning outcomes than any other factor, and both factors (being on task to a large extent and being not at all distracted) were significantly related to greater development in all 4 learning outcomes. Figure 2 shows the difference in months between children who were on task to a large extent and the minority who were not at all on task was equivalent to around 9 months in emergent numeracy and emotion identification, 10 months in emergent literacy and 14 months in mental flexibility (with effect sizes of 0.74 for emergent literacy, 0.84 for emergent numeracy, 0.80 for mental flexibility and 0.60 for emotion identification).

Teachers rated each child's persistence, that is, the extent to which the child usually continued his/her planned course of action in spite of difficulty or obstacles. The difference between the children who are always persistent compared with those who are never persistent was equivalent to around 7 months in emergent literacy, emergent numeracy and mental flexibility, and 8 months in emotion identification (with effect sizes of 0.54, 0.64, 0.42 and 0.51 respectively).

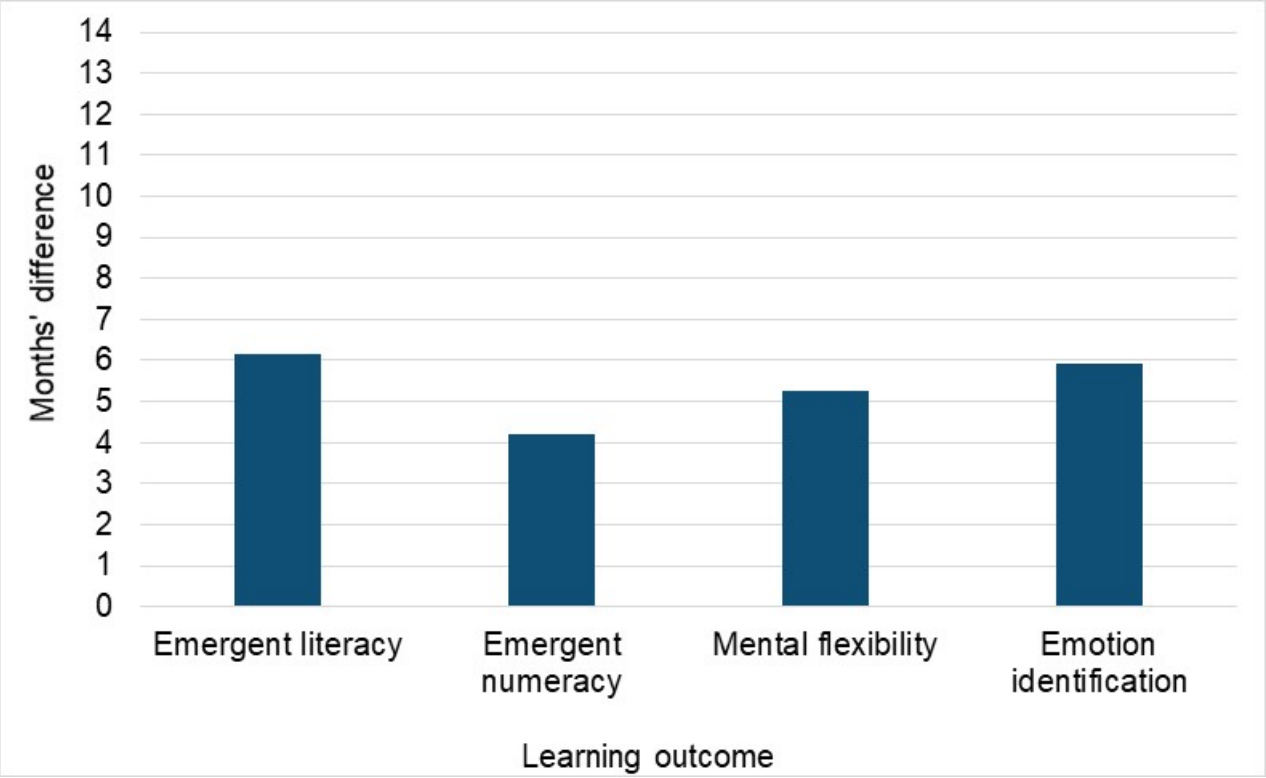
On-task behaviour and persistence had a stronger relationship with children's development than measures of deprivation at age 5. Other factors which were largely protective across learning outcomes were: being in Year 1, rather than in Reception; being older at the time of the assessment; and parental involvement in activities at their child's school (such as school fetes, concerts and parents' evenings).

Children's persistence may be considered malleable and something which can be promoted in early years and school settings. For example, Leonard and others (2019) found that adults are able to influence children's persistence with their words and actions. Persistence appears related to children's executive function (that is, their higher-order cognitive ability), which may be open to influence through classroom-based pedagogy and specific interventions (Diamond, 2012; Serpel and Esposito, 2016; Ackerman and Friedman-Kraus, 2017; Schmidt and others, 2017; EEF, 2021). These findings suggest that further work is warranted to identify effective strategies to increase children's persistence. It is important for teachers to identify children who struggle with attention and mental flexibility because these issues may be responsive to intervention. They can also be indicators of undiagnosed conditions such as Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder.

Having a special education need is a risk factor for all 4 learning outcomes

Having an identified SEN at the age of 5 was a risk factor for children's development, even when other factors (including deprivation) are considered. It was the only risk factor which was consistently identified across all 4 learning outcomes and the size of these effects were greater than those associated with deprivation.

Figure 3: Months' difference in 4 learning outcomes between children without an identified SEN and those with



Source: IELS assessment of 2,577 children, age 5

Children identified as having a SEN had lower development equivalent to approximately 6 months in emergent literacy and emotion identification, 5 months in mental flexibility and 4 months in emergent numeracy (with effect sizes of 0.46, 0.40, 0.31 and 0.41 respectively). This is consistent with existing evidence, for example, the 2019 EYFSP results (DfE, 2019a) found children with an identified SEN were less likely to have reached at least the expected standard in all Early Learning Goals.

Although a fixed characteristic, the analysis demonstrates the importance of identifying children with SEN as early as possible and supporting their development. This is especially important at this time, given that these children are likely to have been disproportionately affected by the Covid-19 pandemic (Pascal and others, 2020).

Some factors were significantly associated with just one particular learning outcome. For example:

- parents having regular back-and-forth conversations with their children about their feelings was positively related to emotion identification development. Children whose parents had these conversations 5-7 days in a week had greater development than those whose parents never had these conversations with their children, equivalent to around 6 months' difference

- being a girl was also a protective factor for emotion identification (a difference of around 4 months, compared with being a boy)
- children with EAL had lower development in emergent literacy than those who had English as their first language (around 4 months' difference).

Are the protective and risk factors different for those children living in deprived areas?

Most of the protective and risk factors identified for all children (namely being older/in Year 1, being on task/not distracted and having an identified SEN) were the same for children living in the most and least deprived areas (comparing children living in the most and least deprived child-level IDACI quartiles) for at least one learning outcome. As these protective factors were largely consistent across models, it suggests these are common to all children, regardless of their socio-economic backgrounds.

The following risk and protective factors were specific to children living in the most deprived areas.

- Having more than 50 children's books in the home (compared with having no children's books in the home) was associated with greater development in emergent literacy and emergent numeracy for children living in the most deprived areas.
- Persistence was associated with emergent numeracy development among children in the most deprived areas but not for those living in the least deprived areas.
- Having EAL was associated with lower emergent literacy development among children living in the most deprived areas.

Conclusion

Both family and school deprivation are related to lower development in emergent literacy, emergent numeracy, mental flexibility and emotion identification at age 5

Both family and school-level deprivation were associated with lower development among children in the IELS sample. School-level IDACI and the parental SES index appear to be the most useful measures to capture the relationship between deprivation and early cognitive outcomes. Measures of parental job role and educational level add further explanatory power to purely income-related measures of deprivation. School-level IDACI is more strongly associated with young children's cognitive development than school-level FSM or parental SES. There are a range of policies and programmes aiming to

address the development gap between disadvantaged children and their peers. These insights can help to ensure that they are targeted effectively.

This study is timely, as the national lockdowns due to the Covid-19 pandemic are likely to have disproportionally affected the most vulnerable children, including those living in poverty, and children with SEN or EAL (Pascal and others, 2020; Ofsted 2020).

A range of risk and protective factors are related to young children's development at age 5

Young children's learning outcomes are influenced by a wide range of factors, of which deprivation is just one. Children's ability to stay on task and persistence had the strongest positive relationships with all 4 learning outcomes. Further investigation is merited into the relationship of these behaviours with executive function and the effectiveness of interventions designed to improve them.

An increased awareness and understanding of the risk and protective factors for young children's development is beneficial for families, ECEC practitioners, teachers and policy makers. Particular focus should be given to the multiple factors related to the development of children from deprived backgrounds, and the role of persistence, the availability of children's books in the home, and the language development of children with EAL as specific risk factors for the development of children who live in deprived areas.

Glossary

Attention deficit-hyperactivity disorder (ADHD) – a cognitive condition characterised by persistent symptoms of inattention and/or hyperactivity-impulsivity (Piek and others, 1999).

Autistic spectrum disorders (ASD) – a range of disorders comprising 3 subcategories: autism, pervasive developmental disorder, and Asperger's syndrome, which are categorised by a range of social and communication impairments and repetitive behaviours (Bhat and others, 2011).

Early Years Foundation Stage profile (EYFSP) – summarises and describes children's attainment at the end of Reception Year. Children's level of development is assessed against the early learning goals (ELGs) and practitioners indicate whether children are meeting expected levels of development, exceeding them or not yet reaching expected levels.

Effect size (ES) – a statistic showing the magnitude of a relationship between 2 variables (one of which is the dependant variable) in a population, taking account of the spread of the distribution. It allows comparisons between variables measured via different scales.

Emergent literacy – an IELS tablet-based assessment focused on 3 areas of language and literacy: listening comprehension, vocabulary knowledge, and phonological awareness.

Emergent numeracy – an IELS tablet-based assessment defined as the ability to recognise numbers and to undertake numerical operations and reasoning in mathematics. The measure focused on simple problem-solving and the application of concepts and reasoning in: numbers and counting, working with numbers, shape and space, measurement, and pattern.

Emotion identification – an IELS tablet-based assessment within the social and emotional domain designed to capture children's ability to identify others' emotional states.

Executive function – the higher order cognitive abilities and processes such as working memory, mental flexibility and self-control that enable people to plan, focus attention, remember instructions, and work on multiple tasks.

Fine motor skills – the ability to use the smaller muscles of the hands to achieve small-scale movements, commonly in activities like using pencils and scissors.

Free school meals (FSM) – a measure of economic disadvantage based upon a child's eligibility to be in receipt of free school meals.

Gross motor skills – the ability to use of the large muscles of the body for walking, running, sitting, jumping and other activities.

Home learning environment – the combination of both the physical characteristics of the home and the quality of the implicit and explicit learning support children receive from parents[3].

IELS – International Early Learning and Child Well-being Study.

Imputation – the statistical process of replacing missing data with substituted values based on other available information, with to the aim of creating a complete dataset.

Income Deprivation Affecting Children Index (IDACI) – an area-level measure of socio-economic disadvantage. IDACI uses information from the Census to measure the proportion of the population in areas experiencing deprivation relating to low income.

Inhibition – a tablet-based assessment within the self-regulation domain of a child's ability to inhibit an impulsive response in favour of an alternative response.

Low birthweight – identified as being less than 2.5kg at birth.

Mental flexibility – a tablet-based IELS assessment within the self-regulation domain focused on a child's ability to shift between rules according to changing circumstances or to apply different rules in different settings.

National Pupil Database (NPD) – a longitudinal database of all children in maintained schools in England. The NPD is compiled and controlled by the Department for Education (DfE) and contains data from a number of distinct datasets. The NPD includes data on pupil and school-level characteristics (such as age, gender, ethnicity, attendance, eligibility for free school meals) linked to data on national curriculum tests and public examinations results.

Persistence – a rating of the extent to which a child continues his/her planned course of action in spite of difficulty or obstacles.

Self-regulation – characterised by a child's ability to think before acting, persist at an activity, follow directions, remain calm, and control their impulses. In IELS, the self-regulation domain focused on 3 distinct measurements: inhibition, working memory and mental flexibility. These are primarily measures of children's cognitive function

[3] Throughout this report, the term 'parents' is used to refer to children's parents and carers.

(sometimes called ‘executive function’) rather than measures of behavioural self-regulation.

Social-emotional learning – a child’s ability to begin forming positive relationships with others, to understand and develop behavioural expectations for both themselves and others, and to understand appropriate behaviour in different settings. IELS measured 5 aspects of children’s social-emotional development, namely: emotion identification; emotion attribution; prosocial behaviour; trust; and non-disruptive behaviour.

Socio-economic status (SES) – a parental SES index derived from responses given in the parent questionnaire relating to parents’ level of education, income and type of employment (OECD, 2020b).

Working memory – a tablet-based assessment within the self-regulation domain focused on a child’s ability to store information and manipulate it to complete a given task.

1 Introduction

The International Early Learning and Child Well-being Study (IELS) is a new study conducted by the OECD. It seeks to understand children's abilities at age 5, and how these are influenced by children's early education experiences, the home learning environment (HLE), and individual and family characteristics[4]. IELS measured the development of almost 7,000 5-year-olds across 3 OECD countries: England, Estonia and the United States. In England, the IELS fieldwork was conducted from October to December 2018, with a nationally representative sample of 2,577 children from 191 schools. The study achieved a high response rate in England, with 95% of the sampled schools and 92% of sampled children from these schools taking part.

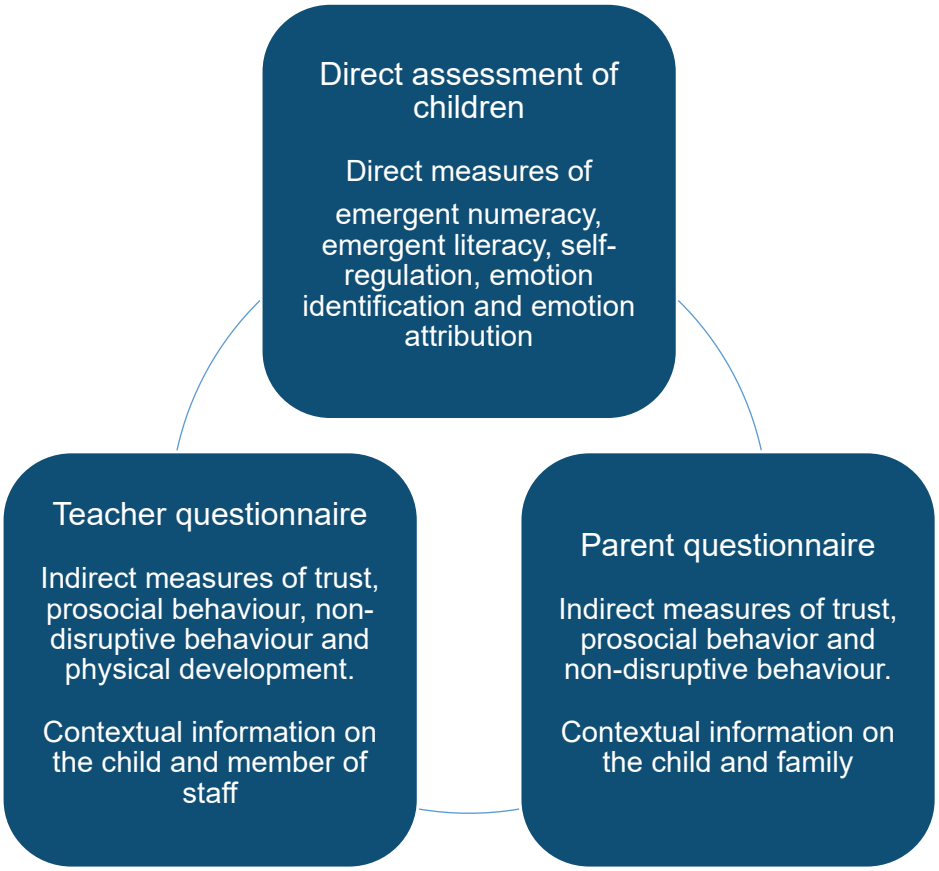
IELS measured children's development in emergent literacy, emergent numeracy, self-regulation and social-emotional development. In England, a teacher assessed module on physical development was added to IELS.

Children were assessed directly by undertaking games and activities on a tablet, supported by a trained and experienced study administrator. IELS also assessed children indirectly, using questionnaires completed by parents and teachers[5]. Most early learning outcomes were assessed both directly and indirectly.

The diagram below outlines the data collection methods used.

[4] Note that although the majority of the children were aged 5, the sample also included some younger children who were aged 4 years 11 months and some older children who were aged 6 years 0 months at the time of assessment.
[5] Around 67% of parents completed the parent questionnaire and around 90% of teachers completed the teacher questionnaire. For full details, see Appendix A.

Figure 4: The different elements of IELS data collection



What is the focus of this report?

Understanding the influence of deprivation on children's development and its interplay with other demographic factors is of particular importance in tackling the early years disadvantage gap. By the time children are aged 5 and have started school, there is already an attainment gap between deprived children and their less deprived peers, equivalent to 4.3 months of learning (Education Select Committee, 2016). The deprivation gap has not closed over the last few years (Education Select Committee, 2016; DfE, 2018a; Hutchinson and others, 2018; Children's Commissioner, 2020; Hutchinson and others, 2020). In fact, it is likely to have widened recently due to the impacts of the Covid-19 pandemic, with deprived children being likely to have fallen further behind in their curriculum learning than their peers (Rose and others, 2021). It is therefore important to identify the factors that are statistically significantly associated with lower development for deprived children and the other factors that are related to all young children's development at age 5. This will inform policy aiming to reduce the deprivation gap.

IELS provided an opportunity to look in detail at how deprivation relates to young children's development in both cognitive and non-cognitive domains. This report expands

on the IELS national findings for England (Kettlewell and others, 2020b) by looking in more depth at the development of children from a deprived background, the factors that influence their development and the protective and risk factors affecting children's development at age 5.

This report will explore the following research questions.

1. How do different measures of socio-economic deprivation and school deprivation interact with each other to predict early learning outcomes at age 5, as measured in IELS?
2. What are the protective and risk factors that can be identified in relation to early learning outcomes at age 5 for all children? Are these different for those children from more deprived backgrounds?

A note on statistical prediction

In relation to statistical analysis, 'prediction' means the extent to which there is an association between an independent variable and a learning outcome which is unlikely to have occurred by chance. This report uses multivariate multi-level models, which take account of the influence of other predictor variables and the inherent structure of the data whereby pupils are 'nested' within schools. The predicted value of each variable represents the strength of an association once the influence of other variables in the model have been taken into account. However, it does not mean that the association is necessarily causal, nor should it be assumed that the results can be used to predict an association between a variable and an outcome at a later point in time.

Chapter 2 considers the different measures of deprivation (national and international measures) and investigates which ones are significantly associated with development at age 5. Chapter 3 presents which factors (children's demographics and their personal and home characteristics) are protective or risk factors for all 5-year-olds' development and considers whether the same characteristics are also risk and protective factors for children from the most and least deprived backgrounds.

Methodology

A similar analytical approach was used for both research questions. The analysis primarily used linear multi-level modelling to test whether independent variables (such as children's age, gender and aspects of the home learning environment) were significant predictors of the learning outcomes (listed below). Multi-level modelling was used to account for the clustering of children within schools. A two-stage process was used for

each model to identify the independent variables that significantly predicted the learning outcomes individually and then improved the predictive ability of the final model. As a result, the analysis identified which variables had the strongest predictive value and modelled their relative effects taking other variables into account. Further details on the methodology are available in Appendix A.

For all models, 4 categories of independent variables were included: National pupil database (NPD) variables; child dataset variables; parent questionnaire variables; and teacher questionnaire variables. However, please note that, due to the need to avoid identifying individual research participants, any analysis resulting in cell sizes of less than 10 has been supressed and the related findings are not described in this report. To comply with rules on avoiding disclosure, additional cells may have been suppressed if their inclusion resulted in the calculation of the supressed cell. Further details on the variables included in each model can be found in Appendix C.

The analysis focused on 4 of the IELS learning outcomes. These were all direct measures of children's development, rather than indirect measures assessed by teachers.

- **Emergent literacy**[6] and **emergent numeracy** – chosen for their well-documented importance for children's development and education policy
- **Mental flexibility** – a self-regulation measure of children's ability to shift their thinking according to the circumstances; chosen due to its relationship with free school meals (FSM) already identified and its strong correlation with the other chosen learning outcomes
- **Emotion identification** – a social emotional measure of children's ability to identify other people's emotions; chosen due to its relationship with FSM already identified and its strong correlation with the other learning outcomes.

These 4 learning outcomes were chosen to provide a breadth of measures across different domains. All 4 learning outcomes included in this report are also explored in a second thematic report using IELS data (Lucas and others, 2021), which looks in more detail at children's physical development.

Prior to conducting the analysis, the research team needed to address the issue of missing data. Aside from the learning outcome measures, data for all variables required imputation. This particularly affected data from the parent questionnaire – (such as parental socio-economic status (SES), number of siblings, low birthweight and the home learning environment) because 33% of parents of participating children did not complete the parent questionnaire. For details on the imputation methodology see Appendix A.

[6] Note that emergent literacy focused on oral language rather than reading or writing.

What is already known about deprivation and early development?

Young children's development in both early literacy and numeracy shows evidence of a gap in development associated with the SES of their parents. The 2018 EYFSP results (DfE, 2018a) show a difference of 10 percentage points in communication and language development between children who were eligible for free school meals (FSM) and those who were not. The equivalent gap for mathematics was 17 percentage points.

Previous research also suggests that children from households with a lower SES show lower development in executive function and cognitive self-regulation (McClelland and others 2015, Ackerman and Friedman-Krauss, 2017). Furthermore, young children from lower-income households tend to exhibit less advanced social-emotional development, to exhibit more disruptive behaviour and have poorer emotional health (Feinstein, 2015; Chowdry and McBride, 2017). These differences are evident by the age of 3 (Feinstein, 2015). Chowdry and McBride (2017) calculated that, at age 5, the gap between the highest and lowest socio-economic class groups in terms of disruptive behaviour and emotional health was 0.6 standard deviations but the difference was smaller for children who were older.

With regards to early childhood education and care (ECEC), the Study of Early Education and Development (SEED) found that by the start of school almost all children, regardless of deprivation, had used some form of formal group ECEC (Melhuish and Gardiner, 2020). For the 40% most deprived children, starting to use a minimum of 10 hours per week of formal ECEC no later than age 2, combined with a mean use of over 20 hours per week of formal ECEC between age 2 and the start of school, increases the chances of achieving expected EYFSP levels in school Reception year and improves children's verbal ability in school Year 1. Although childminder use is a less common form of ECEC, the SEED study also found that for children from the 20% most deprived group, childminder use was associated with a lower total of EYFSP score, though this relationship was not seen across groups who were less deprived.

The national report for IELS (Kettlewell and others, 2020b) showed that children eligible for FSM had significantly lower development across all IELS measures with the exception of inhibition, than their peers who were not eligible for FSM.

This report is timely, due to the concerns about the impact of the Covid-19 pandemic and the risk that this crisis will open up further gaps in development between children from deprived backgrounds and their peers (EPI, 2020; Montacute, 2020). As well as affecting their educational achievement, this is likely to have a negative effect on their future chances in the labour market (Hupkau and others, 2020).

Although there is a body of research into the influence of deprivation on children's education, IELS presents a unique opportunity to explore the relationship between deprivation and development at the age of 5 using the wide range of learning outcomes included in IELS and the wealth of background information collected on each child. As well as focusing on cognitive development, this study adds to the emerging body of research on young children's self-regulation and social and emotional development.

2 How do different measures of deprivation interact to predict learning outcomes at age 5?

Chapter summary

When investigating each of the 5 measures of deprivation individually (namely, child-level FSM, school-level FSM, child-level IDACI, school-level IDACI and parental SES), the analysis showed that children who are more deprived on any measure have lower development than their non-deprived peers in emergent literacy, emergent numeracy, mental flexibility and emotion identification.

However, when all variables were considered together in the multivariate analysis, only one or two of the deprivation measures were related to 5-year-olds' development in each of the learning outcomes.

The multivariate analysis showed that:

- parental socio-economic status (SES) (a combined measure of parental income, occupation and education) was associated with children's development in emergent literacy, emergent numeracy and mental flexibility. Children who had a higher parental SES showed greater development in these areas even when all other factors in the model were considered
- school-level Income Deprivation Affecting Children Index (IDACI; a measure of the deprivation of the area in which the school is located) was associated with development in emergent numeracy and emergent literacy, even when controlling for parental SES. Children who attended a school in a more affluent area showed greater development than those who attended a school in a deprived area
- attending a school with a lower proportion of children eligible for Free School Meals (FSM; a school-level measure of deprivation) was associated with greater development in emotion identification
- child-level FSM was not significantly related to children's learning outcomes in the final models for emergent literacy, emergent numeracy, mental flexibility or emotion identification. This suggests that parental SES (a measure which combined parental income, occupation and education level that was present in 3 of the models) may have a stronger relationship with children's learning outcomes than FSM eligibility alone
- for emergent literacy and emergent numeracy, where both school-level IDACI and parental SES were significantly related to children's learning outcomes, school-level IDACI had a stronger relationship with the outcome measure than parental

SES. School-level IDACI and parental SES operate in an additive fashion without one changing the effect of the other

- Although none of the deprivation measures completely overlapped, 71.3% of children attending a school in an area of high deprivation also attended a school with a high proportion of children eligible for FSM and 71.2% of the children who were eligible for FSM were in the most deprived quartile of children according to the IELS measure of parental SES. This suggests that where children were deprived on one measure, there was a high probability they were also deprived according to at least one other measure.

Introduction

This chapter investigates the relationship between children's development and 5 measures of deprivation available for the IELS sample linked to the national pupil database[7]: free school meal (FSM) eligibility; child Income Deprivation Affecting Children Index (IDACI); school-level IDACI, school-level FSM; and the parental socio-economic status (SES) measure.

Measures of deprivation

There are different ways to measure socio-economic deprivation and several measures are already commonly used for this purpose in England. This analysis sets out to identify which measure or set of measures is best able to capture the relationship between deprivation and children's development so that policies and interventions can be effectively targeted towards those most in need.

Free school meal (FSM) eligibility

Eligibility for free school meals (FSM)[8] is widely used as a proxy for socio-economic status in educational research and policy. Children are eligible for free school meals if they or their family are in receipt of certain benefits and financial support[9]. This measure categorises children into two groups; those 'eligible' and those 'not eligible' for FSM. The measure cannot capture variations within the two categories, but this is not necessarily

[7] The national pupil database is a longitudinal database of all children in maintained schools in England
[8] For the purpose of this investigation, FSM was chosen over pupil premium or Ever FSM. FSM was selected rather than pupil premium, because the pupil premium includes other groups such as looked after children and children of service families. For young children, FSM and Ever-FSM are largely the same, as children have not had time to move in and out of FSM eligibility. Analysis of the dataset confirmed this, as FSM and Ever-FSM were identical in the IELS sample
[9] The eligibility criteria for FSM can be found here: Apply for free school meals

an issue if the aim is simply to determine just who are the poorest in society (Gorard and Huat See, 2013).

A number of researchers have compared FSM with other measures of deprivation and concluded that it is satisfactory for determining educational disadvantage (see Crawford and Greaves, 2013 and Sutherland and others, 2015). However, some have questioned the suitability of FSM eligibility for identifying those most in need for support for two main reasons. Firstly, it is debatable whether the financial threshold for FSM truly captures all children living in poverty. Taylor (2018) identified a small but significant group of children in the Millennium Cohort Study who were living in poverty but were not categorised as eligible for FSM. Similarly, Hobbs and Vignoles (2009) argue that some children are living in 'working poor' households, where their parents earn just above the income threshold and do not or cannot claim other benefits. Conversely, families who do apply for FSM are also likely to be receiving other welfare benefits and may not be among the lower income group once all support is received (Hobbs and Vignoles, 2009).

Secondly, FSM eligibility is reliant on parents applying for the support. In England, universal infant free school meals (UIFSM) are provided for all children from Reception to Year 2, which removes a primary motive of registering for FSM (Holford and Rabe, 2020). However, the findings on the direct impact of UIFSM on FSM take-up within these year groups is currently inconclusive. The rate for registering for FSM after the introduction of UIFSM was found to be lower than usual in one study (Holford and Rabe, 2020) but in another, school leaders differed in their opinions about its impact on take-up (Sellen and others, 2018).

Income Deprivation Affecting Children Index (IDACI)

The Income Deprivation Affecting Children Index (IDACI) is an area-level measure of socio-economic deprivation. IDACI is derived from the proportion of the population in an area experiencing deprivation relating to low income. Schools and household postcodes are given a score based on the Lower Super Output Area (LSOA). There are 32,482 LSOAs in England, each containing about 625 households with a mean population of around 1,500 (HM Government, n.d.). Child IDACI is a continuous measure and aims to capture the wider circumstance in which children are living in.

Critics of this measure point out that it can mask individual circumstances, as some of the most deprived families live in areas which are not classified as deprived (Gorard and Huat See, 2013); and even if areas are similar in terms of population wealth they may differ in other respects, such as access to social services, schools and other educational opportunities (Ilie and others, 2017).

Parental socio-economic status (SES)

IELS included a measure of socio-economic status (the SES index) based on self-reported parental education and parental occupation[10] as well as household income for children in the study. This measure, developed by the OECD, captures more than just economic deprivation by including aspects of social deprivation and cultural capital too.

Social factors within the family are known to influence the impact of deprivation on attainment. Parental education and occupation are examples of cultural capital (Bourdieu and Passeron, 1990). A similar measure to the parental SES in IELS, known as the economic, social and cultural index (ESCS), is used in the Programme for International Student Assessment (PISA)[11]. Wheater and others (2016) reviewed the utility of the ESCS in analysing the PISA results for England. They concluded that more sophisticated measures that address social and cultural deprivation as well as economic deprivation are better at identifying deprivation than just measures of economic deprivation, such as FSM, alone. Additionally, Taylor (2018) found socio-economic factors such as the educational levels of parents appeared to be associated with attainment at age 7. Ilie and others (2017) concluded that parental occupation levels and parental education are the best predictors of pupils' attainment and suggest using parental occupation/education instead of FSM eligibility, although they recognised that this data is less readily available than FSM.

Table 1 summarises the different measures of deprivation investigated in this report.

Table 1: Measures of deprivation in this analysis

Measure	Description	Variable type	Level of measure	Data source
Chid-level FSM eligibility	Child-level eligibility for being in receipt of free school meals. FSM is a binary measure (eligible and not eligible), based whether the family is in receipt of certain benefits.	Binary variable	Child-level measure	National Pupil Database (NPD)

[10] Where the education and occupation levels were available for both parents, the highest levels were used for the analysis.
[11] An international assessment of 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges.

Measure	Description	Variable type	Level of measure	Data source
School-level FSM	The proportion of pupils eligible for free school meals in a school.	Ordinal variable	School-level measure	Government website - Compare school performance
Child[12] IDACI	Each child's home postcode is allocated an IDACI score, representing the proportion of all children in the area aged 0 to 15 living in income-deprived families. The definition of low income includes both those people that are out-of-work, and those that are in work but who have low earnings (and who satisfy the respective means tests).	Continuous variable	Area-level measure included in the statistical models as a child-level measure	NPD
School-level IDACI	The level of deprivation (measured by IDACI as outlined above) in the Lower Super Output Area where each school is located.	Ordinal variable	Area-level measure included in the statistical models as a school-level measure	Government website - Compare school performance
SES Index	Derived from responses to the IELS parent questionnaire on parental occupation, parents' level of education and household income.	Continuous variable	Child-level measure	IELS dataset

[12] This variable is labelled as 'pupil IDACI' in the NPD.

The relationship between deprivation and children's outcomes in IELS

As shown in Table 1, the 5 deprivation measures were either focused on children (and their families) or on their schools. The 3 child-level measures of deprivation are discussed first, followed by the two school-level measures.

The relationship between FSM and children's outcomes at age 5

Information on a child's eligibility for FSM was available for 2,462 (96%) of the children[13] who took part in IELS (Kettlewell and others, 2020b)[14]. In total, 429 children were eligible for FSM (17% of the sample for whom data was available)[15] and 2,033 children (83%) were not eligible for FSM. This is comparable with the national average of 15.4% of primary school children eligible for FSM, based on data collected in January 2019, shortly after the IELS study took place (DfE, 2019c).

Previous analysis of the IELS data in England by Kettlewell and others (2020b) investigated the relationship between FSM and children's learning outcomes using bivariate analysis (that is, looking at the relationship between a child's FSM eligibility and their outcome on each measure, without taking account of the influence of any other variables). They found significant differences between children eligible for FSM and those who were not in all 4 learning outcomes investigated in this report (emergent literacy, emergent numeracy, mental flexibility and emotion identification). Further details are given in Appendix B

The relationship between child-level IDACI and children's outcomes at age 5

A child-level IDACI score was available for 2,454 children in the IELS sample, once matched to the NPD. This equated to 95% of the sample. Based on bivariate analysis conducted for this report[16], there was evidence of a significant relationship between child-level IDACI quartile[17] and all 4 learning outcomes investigated in this report. Further details are given in Appendix B.

[13] Some percentages have been rounded to the nearest whole number here and elsewhere in the report.
[14] Throughout this report percentages taken from the national report (Kettlewell and others (2020b)) may vary marginally to those from this current analysis due to the process of imputing missing values.
[15] Note that the IELS sample was stratified by the percentage of children in the school eligible for FSM as well as school type (local authority maintained, academy or independent) and region.
[16] The bivariate analysis for child-level IDACI is based on the sample of children for whom NPD data was available without imputing missing values (2,454).
[17] In this bivariate analysis, the distribution of child-level IDACI scores was divided into 4 groups or quartiles, with each group containing almost equal numbers of children. However, for the purposes of the multi-level models (detailed later in this chapter and Chapter 3), child-level IDACI was entered as a continuous variable.

The relationship between parental SES and children's outcomes at age 5

Parental SES comprised 3 dimensions derived from responses to the parent questionnaire: parent occupation, parent educational level and household income. The parental SES measure has the benefit of being internationally comparable. OECD (2019a) found that higher parental SES was associated with greater development in emergent literacy, emergent numeracy and emotion identification learning outcomes in all 3 countries which participated in IELS. Higher parental SES was also positively associated with mental flexibility development in England and the United States.

The relationship between school FSM and children's outcomes at age 5

One way to investigate the relationship between school-level deprivation and children's development is to look at the proportion of children eligible for FSM in the school. In line with this, the 191 participating schools were divided into 4 groups related to the national distribution of FSM. This bivariate analysis[18] showed children who attended schools in the highest quartile for FSM had significantly lower development than those in the lowest quartile for school FSM in all 4 measures. Further details are given in Appendix B.

The relationship between school-level IDACI and children's outcomes at age 5

School-level IDACI is a measure of the deprivation of the area in which the school is located. The bivariate analysis[19] found that children who attended schools in the most deprived school-level IDACI quartile had significantly lower development in all 4 measures compared to children who attended schools in the least deprived school-level IDACI quartile. Further details are given in Appendix B.

How do different measures of deprivation interact to predict learning outcomes at age 5?

Overview of the statistical model

As outlined in Chapter 1, this analysis of deprivation primarily used linear multi-level modelling to examine the association between different variables and children's development in emergent literacy, emergent numeracy, mental flexibility and emotion

[18] The bivariate analysis for school-level FSM is based on the sample of children for whom NPD data was available without imputing missing values (2,375).
[19] The bivariate analysis for school-level IDACI is based on the sample of children for whom NPD data was available without imputing missing values (2,520)

identification. This included 5 measures of deprivation as described in Table 1. To be included in the final model, the independent variables had to significantly predict the dependent variable (that is, the learning outcome) individually and then had to improve the predictive ability of the multi-level model over and above the other variables in the model (see the Appendix A for further details of the methodology).

The final model included measures of deprivation, children's demographic characteristics and other home and background characteristics. The findings presented here discuss the deprivation measures which were included in the final model for each of the learning outcomes in order to address the first research question which focused on the interaction between measures of individual and school-level deprivation. The other characteristics (related to the second research question), are discussed in Chapter 3. Table 2 shows which measures of deprivation were significantly related to 5-year-olds' development.

Throughout this report, only factors that were significantly related to children's development in at least one learning outcome are reported; and any differences reported are statistically significant at the $p<0.05$ level. Where a factor is significantly associated with a learning outcome the effect size is given in the accompanying text. Effect size shows the magnitude of a relationship between 2 variables in a population, accounting for its distribution. The standardised coefficients for each factor included in the models are given in Appendix C.

To add further context to the findings in this report, where possible, differences between groups have been translated into differences in months of development[20]. On IELS measures where there were significant differences by age of child, the difference was converted into a monthly difference value, which provides the number of points gained with each additional month of age. This was then used to calculate the approximate difference in months between the scores of two groups (for example, girls and boys), to quantify, in relative terms, how far ahead or behind one group was of the other[21]. Further information about this calculation can be found in Appendix A.

Deprivation measures significantly associated with 5-year-olds' development

Table 2 shows which of the 5 deprivation measures were significantly related to each of the 4 child learning outcomes in the multivariate models (full details of all the multi-level models are available in Appendix C).

[20] Please note that the findings have not been age-standardised.
[21] Note that the estimate of the difference by age in months reported here may differ from that reported by the OECD (2020b) because this calculation includes children aged 4 years and 11 months, which were excluded from the OECD's analysis.

Table 2: Deprivation measures significantly associated with 5-year-olds' development

Emergent literacy	Emergent numeracy	Mental flexibility	Emotion identification
School-level IDACI*	School-level IDACI*	SES index*	School FSM*
SES index*	SES index*	-	-

* This is statistically significant ($p < 0.05$).

Source: IELS assessment of 2,577 children, age 5

Child-level measures of deprivation

A higher score on the parental SES measure (a continuous measure) was associated with greater development in emergent literacy, emergent numeracy and mental flexibility. The effect sizes were relatively small at 0.13 for emergent literacy, 0.14 for emergent numeracy and 0.09 for mental flexibility, equivalent to about 2 months' difference for emergent literacy and mental flexibility and around 1 month's difference for emergent numeracy[22]. Nevertheless, this does suggest that parental income, occupation and parental education may be important for young children's development in all 3 of these learning outcomes.

The absence of significant relationships between the learning outcomes and child-level FSM suggests that its inclusion did not add any more information once an alternative measure of deprivation was already included in the model. For example, once school-level IDACI and the parental SES measures were included in the emergent numeracy model, no further measures increased the predictive power of the model. This is likely to be because the deprivation variables were correlated with one another[23]. For example, the correlation between child-level IDACI and FSM was 0.32 and the correlation between parental SES and child-level FSM was 0.28. Additionally, over half (51%) of pupils eligible for FSM were in the most deprived school-level IDACI quartile compared with 9% who were in the least deprived school-level IDACI quartile. There was a similar pattern of relationships between child- and school-level FSM: over half (54%) of pupils eligible for

[22] As SES is a continuous variable, the effect size calculated represents the effect of a change in the independent variable (that is, SES score) equivalent to its adjusted standard deviation on the outcome of interest.

[23] Two checks on the multicollinearity of variables were conducted to ensure that variables entered into the models were independent of one another. This avoids entering two variables which are highly correlated, which can result in significant relationships being missed. These checks used a threshold of above 0.9 for Pearson's correlation and above 10 using Variance Information Function. Further details on the assumption checks are given in Appendix A.

FSM were in the highest quartile for school FSM, compared with just 4% in the lowest quartile.

School-level measures of deprivation

For both emergent literacy and emergent numeracy, being in the least deprived and second least deprived quartiles for school-level IDACI was significantly associated with greater development compared with attending schools located in the most deprived areas. There was a difference of approximately 5 months in emergent literacy and 4 months in emergent numeracy between children attending school in the least deprived areas compared with the most deprived areas. The effect sizes were 0.35 for emergent literacy and 0.38 for emergent numeracy, indicating that the deprivation of the area surrounding a child's school may have an effect on their emergent literacy and emergent numeracy development at age 5. These effect sizes were comparatively larger than the effect sizes for parental SES, suggesting school-level IDACI predicts more variance in these learning outcomes than SES. Few studies account for the complex interplay between individual and area whilst accounting for socio-economic status, particularly for children in England (Crenna-Jennings, 2018). Some studies suggest that community networks and community interaction as well as access to area-level resources, such as green space and after-school programmes, are important for children's development (see Crenna-Jennings, 2018).

Emotion identification was only significantly related to one deprivation measure: school-level FSM. Children attending schools where the proportion of FSM pupils is higher (that is, schools in the highest FSM quartile compared to those in the lowest FSM quartile) had lower development in emotion identification. The effect size was 0.24, which is equivalent to about 4 months' difference in emotion identification. The proportion of FSM pupils in a school is likely to be related to area-level deprivation (that is, school-level IDACI), but in this case it was the proportion of deprived pupils attending the school, rather than the deprivation of the school location which contributed to the model. This a unique finding compared to the other learning outcomes, and may suggest that there is a specific association between the demographics of the school intake and young children's development in emotion identification. This possibility is supported by findings from the Effective Pre-school and Primary Education 3-11 Project (EPPE 3-1, Sylva and others, 2008) which suggest that the overall characteristics of a school's pupil intake influence children's social outcomes.

The interaction between deprivation measures

In order to investigate the relationship between individual deprivation, area-level deprivation, and children's learning outcomes, a series of interactions were run in the models. Interactions can help to provide the context of a relationship between two predictors by investigating whether the effect of one predictor influences (that is,

changes) the effect of another predictor on the learning outcomes. For example, whether the disadvantage associated with both living in a deprived family and attending a school in a deprived area was straightforwardly cumulative, or whether the two types of disadvantage together intensified (or multiplied) any negative relationship with learning outcomes.

The 4 interactions run in the models were:

- a child-level IDACI by school-level IDACI interaction
- a school-level IDACI by parental SES interaction
- a parental SES interaction with all factors (such as family and child characteristics which are discussed in Chapter 3)
- school-level IDACI interacted with all factors.

None of these interactions led to an increase in the amount of variance explained in the learning outcomes over the main effects models (that is, the models that are presented in this report). For example, the effect of going to school in a deprived area did not intensify the effect of having a low parental SES. So, where school-level IDACI and parental SES are significant main effects in the same model (such as for emergent literacy) they operate in an additive fashion, without one changing the effect of the other. Also, where other factors such as parental involvement (discussed below) are significantly associated with children's development, these influences occur regardless of parental SES and the deprivation of the area in which the school is located.

The overlap between deprivation measures

Child-level FSM, child-level IDACI, school-level IDACI, school-level FSM and parental SES measure deprivation in different ways. Table 3 shows the overlap in the proportion of children considered most deprived across each of the measures. Each row focuses on one measure of deprivation and its overlap with the other 4 measures. The percentages presented are the proportion of children defined as deprived using that measure who would also be defined as deprived using another measure.

As the tables show, there are no measures which completely overlap. Over half of children in the IELS sample who were eligible for FSM also lived in an area of high deprivation, attended a school in an area of high deprivation or attended a school with a high proportion of FSM-eligible children. Over 70% of children who were eligible for FSM were children in the most deprived quartile using the parental SES measure. On the other hand, across the measures, the proportion of deprived children (as determined by each measure) who were also eligible for FSM was less than 40%. This is likely to be reflecting that the other 4 measures (child-level IDACI, school-level IDACI, school-level

FSM and parental SES) are split into quartiles, as opposed to the binary measure of FSM, and therefore FSM children will be more spread across the different quartiles.

The area-level measures of school and child-level IDACI appear to be most closely associated as well as school-level IDACI and school FSM. Over 70% (71.3%) of the children attending a school in an area of high deprivation also attended a school with relatively high proportions of FSM-eligible children. In addition, 60.2% of children who attended a school in a deprived area also lived in a deprived area.

Table 3: Overlap between the deprivation measures

-	Child-level FSM	Child-level IDACI	School-level IDACI	School-level FSM	Parental SES
Children who were eligible for free school meals	-	51.5% also lived in the most deprived child-level IDACI quartile (n=219)	50.7% attended a school in the most deprived school-level IDACI quartile (n=213)	53.7% were in the highest quartile for proportion of FSM-eligible children in their school (n=220)	71.2% were in the most deprived quartile for parental SES (n=148)
Children who lived in the most deprived areas (that is, the most deprived child-level IDACI quartile)	35.7% were eligible for FSM (n=219)	-	67.9% also attended a school in the most deprived school-level IDACI quartile (n=404)	63.9% were in the highest quartile for proportion of FSM-eligible children in their school (n=376)	52.0% were in the most deprived quartile for parental SES (n=178)
Children who attended schools in the most deprived areas (that is, the most deprived school-level	31.6% were eligible for FSM (n=213)	60.2% lived in the most deprived child-level IDACI quartile (n=404)	-	71.3% were in highest quartile for proportion of FSM-eligible children in their school (n=483)	45.7% were in the most deprived quartile for parental SES (n=169)

-	Child-level FSM	Child-level IDACI	School-level IDACI	School-level FSM	Parental SES
IDACI quartile)					
Children who attended schools with the highest proportions of FSM-eligible children (that is, the most deprived quartile)	36.1% were eligible for FSM themselves (n=220)	62.3% lived in the most deprived child-level IDACI quartile (n=376)	79.1% also attended a school in the most deprived school-level IDACI quartile (n=483)	-	48.4% were in the most deprived quartile for parental SES (n=156)
Children in the most deprived quartile for parental SES	36.1% were eligible for FSM (n=148)	43.6% lived in the most deprived child-level IDACI quartile (n=178)	41.7% also attended a school in the most deprived school-level IDACI quartile (n=169)	39.4% were in highest quartile for proportion of FSM-eligible children in their school (n=156)	-

Source: IELS assessment of 2,577 children, age 5

Parental SES and school-level measures of deprivation appear to be the least highly related. This is reflected in the two ways one can read the overlap. Firstly, of the group of children who are in the most deprived quartile for SES, the proportion of children who also attend a school in the most deprived areas is below 50 per cent (41.7%). This was also the case for the proportion of children in the most deprived quartile for parental SES who were also in the highest quartile for the proportion of FSM-eligible children in their school (39.4%). Looking at it the other way, the proportion of children who attend a school in the most deprived areas and are in the lowest quartile for SES is again under 50% (45.7%). This is also true when focusing on the group of children in the highest quartile for the proportion of FSM-eligible children in their school.

3 Which protective and risk factors are related to learning outcomes at age 5?

Chapter summary

The multivariate analysis provided evidence for a variety of protective and risk factors associated with development.

Protective factors

The factors protective of children's development included child behaviour characteristics, child demographics as well as parental engagement with school. This analysis showed that when all other factors (including deprivation) are considered:

- being older (for example, being aged 6 years 0 months compared with 5 years 0 months) was significantly associated with greater development in all 4 learning outcomes of emergent literacy, emergent numeracy, mental flexibility and emotion identification
- being more persistent and being on task/not distracted was also significantly associated with greater development in all 4 learning outcomes of emergent literacy, emergent numeracy, mental flexibility and emotion identification
- being in Year 1 rather than Reception was a protective factor for children's development in 3 learning outcomes: emergent literacy, emergent numeracy and mental flexibility
- having parents who are strongly/moderately involved with activities in their child's school was a protective factor for 3 learning outcomes: emergent literacy, emergent numeracy and emotion identification.

Risk factors

The risk factors associated with children's development were mainly child and family demographics. This analysis showed that when all other factors (including deprivation) are considered:

- having an identified special educational need (SEN) was a significant risk factor for children's development in emergent literacy, emergent numeracy, mental flexibility and emotion identification
- children with English as an additional language (EAL) had lower development in emergent literacy than those with English as their first language.

To investigate further the relationship between deprivation and children's development, risk and protective factors were modelled for children living in the most, compared to the least, deprived areas (measured by child-level IDACI quartile)).

This more in-depth analysis provided further support for some of the protective and risk factors identified in the whole sample (namely being older, being on task/not distracted and having an identified SEN) with limited evidence that these factors affected the most and least deprived children differently. As these protective factors were largely consistent across models, it suggests they are not confined to comparisons between extremes of deprivation.

The following risk and protective factors were specific to children living in the most deprived areas.

- Persistence was a protective factor for emergent numeracy development.
- Having more than 50 children's books in the home was protective of greater development in emergent literacy and emergent numeracy compared with having no books in the home.
- Having English as an additional language was associated with lower emergent literacy development.

Introduction to protective and risk factors

This chapter addresses the research question of what protective and risk factors can be identified for the 4 learning outcomes, by examining other variables (such as family and child characteristics) that may be positively or negatively related to children's outcomes, over and above the influence of deprivation. It identifies which of the individual and family variables available in the IELS dataset (linked to the national pupil database) were found to be related to children's emergent literacy, emergent numeracy, mental flexibility and emotion identification development after taking account of the influence of other variables included in the models, including measures of deprivation. In order to address the second part of the research question, whether the risk and protective factors are different for those children from more deprived backgrounds, this chapter also explores whether these same variables are related to development for children living in the most and least deprived areas.

Overview of the statistical models

As outlined in Chapter 1, linear multi-level modelling was used to examine the association between different variables (most of which were related to children's individual and family characteristics) and each of the 4 learning outcomes.

Each factor that significantly contributed to the models is discussed below in terms of 'protective' and 'risk' factors. As the names suggest, protective factors are those which are associated with greater development and risk factors are those associated with lower development in each of the learning outcomes. Therefore these are characteristics which may warrant further attention in policy and practice.

The protective and risk factors associated with 5-year-olds' development

Table 4 summarises the risk and protective factors across all 4 learning outcomes. It includes all protective and risk factors that were significantly associated with 2 or more of the learning outcomes. The full range of protective and risk factors that were significantly associated with each learning outcome are discussed below.

Table 4: Risk and protective factors associated with 5-year-olds' development in emergent literacy, emergent numeracy, mental flexibility and emotion identification

Protective factors	Risk factors
Attention and persistence: • Children who were on task to a large extent during the direct assessment, compared to those not at all on task (all 4 learning outcomes) • Children who were not at all distracted during the direct assessment, compared to those distracted to a large extent (all 4 learning outcomes) • Children who were always persistent, compared to children who were never persistent (all 4 learning outcomes) Age and school year: • Children who were older at the time of assessment (all 4 learning outcomes) • Children in Year 1 compared to children in Reception (all learning	Child and demographic characteristics: • Children with an identified SEN (all 4 learning outcomes) • Children from backgrounds with lower parental SES (all learning outcomes except emotion identification) • Children who attended a school in the most deprived IDACI quartile, compared with children who attended a school in the least deprived IDACI quartile (emergent literacy and emergent numeracy only)

outcomes except emotion identification) Parental involvement: • Children whose parents were strongly or moderately involved with activities at their child's school, compared to parents who were slightly or not involved (all learning outcomes except mental flexibility)	

Source: IELS assessment of 2,577 children, age 5

Figures 5 to 8 show the mean effect sizes[24] of the factors significantly associated with each learning outcome. Where a factor is significantly associated with a learning outcome, the effect size is given in the accompanying text, together with an estimate of the approximate difference in months (more detail on this calculation can be found in Appendix A).

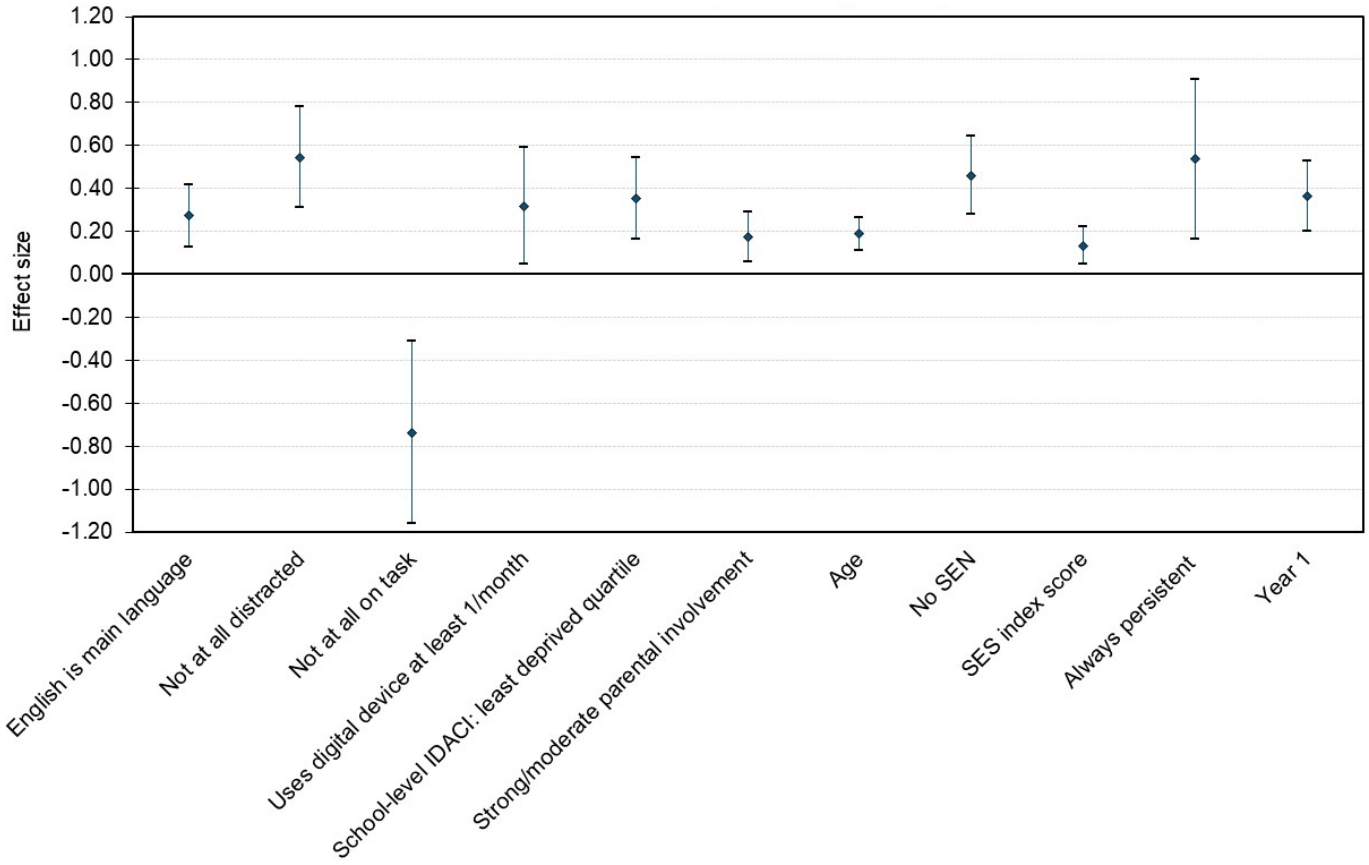
Where multiple categories for the same variable emerged as significant, the category with the largest effect size is shown. The diamonds depict the mean effect size, while the whiskers above and below the mean show the 95% confidence intervals, taking account of the standard error, meaning that there is 95% confidence that the true value lies within this range.

Factors associated with emergent literacy development

Figure 5 shows the factors that were significantly associated with children's literacy development, including deprivation measures along with other family and individual characteristics and some aspects of children's behaviour.

[24] This is the absolute effect size between the variable in the model and its comparator.

Figure 5: Factors associated with emergent literacy development



Source: IELS assessment of 2,577 children, age 5

Figure 5 shows that a range of factors were significantly associated with children's development in emergent literacy. The factors with the strongest relationship (as indicated by the effect size) were on-task behaviour, distraction and persistence.

Emergent literacy – on task, distraction and persistence

IELS collected data on children's attention and distraction. IELS study administrators were asked to rate 2 aspects of children's behaviour during the direct assessments: 'Was the child easily distracted?' and 'Did the child stay on task?'[25]. The study administrators responded via a 4-point scale ('not at all', 'not really', 'to some degree' and 'to a large extent').

The small minority (1%) of children who were not at all on task during the direct assessments had significantly lower emergent literacy development than the 60% who were on task to a large extent. This factor had an effect size of -0.74 which is equivalent

[25] Note that these variables (attention and distraction) appear to be measuring different aspects of children's behaviour as the correlation between the two was not sufficiently high to indicate that only one of them should be included in the statistical models. Please see Appendix A for further details of the assumption checks conducted for the analysis.

to approximately 10 months for emergent literacy. The effect of being not at all on task was associated with larger confidence intervals[26], the upper confidence interval was -0.31 and the lower confidence interval was -1.16. There was also a significant difference between the 60% of children who were on task to a large extent and the 32% who were on task to some degree (see Appendix C).

Two other factors had strong positive associations with emergent literacy, namely being not at all distracted during the direct assessments and always being persistent.

The 42% of children who were rated by study administrators as being not at all distracted during the direct assessments had significantly greater development in emergent literacy than the 7% of children who were rated as being distracted to a large extent. The effect size for this relationship was 0.55, which is equivalent to approximately 7 months' difference in emergent literacy.

Teachers were asked to rate each child's persistence, that is, the extent to which the child continued his/her planned course of action in spite of difficulty or obstacles. Teachers rated 4% of children as never persistent, 15% of children as rarely persistent, 48% of children as sometimes persistent, 27% of children as often persistent and 6% of children as always persistent (Kettlewell and others, 2020b).

The 6% of children who were rated by their teachers as always persistent had significantly greater development in emergent literacy than the 4% who were never persistent. The effect size was 0.54 which is equivalent to approximately 7 months of difference in emergent literacy. The effect of being always persistent was associated with larger confidence intervals, the upper confidence interval was 0.91 and the lower confidence interval was 0.17.

The deprivation measures of school-level IDACI and parental SES had relatively lower effect sizes[27] than distraction, on-task behaviour, persistence, having an identified special educational need (SEN) and being in Year 1.

Emergent literacy – English as an additional language and digital device use

Two factors that were significant in the model for emergent literacy were not identified in the other 3 learning outcomes, namely having English as an additional language (EAL) and using a digital device at home.

Previous research has found that having EAL is associated with lower levels of literacy and numeracy in the early years, but this effect reduces markedly with age and is largely

[26] The confidence interval represents the range of values within which we can be 95% confident the true mean of the population lies

[27] It is worth noting, however that some of the relationship between persistence and attention could relate to deprivation, if children had experienced different deprivation levels before age 5. However, this would be hard to capture in a cross-sectional study of this kind.

eliminated by age 16 (Strand and others, 2016; Gorard and Siddiqui, 2019; DfE, 2019b). The IELS national report for England (Kettlewell and others, 2020b) found that children identified in the national pupil database (NPD) as having EAL showed lower development than their peers in emergent literacy, emergent numeracy and mental flexibility.

The multivariate analysis reported here found that, when all other factors are considered, having EAL was a significant risk factor for children's emergent literacy development, with an effect size of 0.27. The difference between children with EAL (17% of the sample) and their peers was equivalent to 4 months in emergent literacy. There was no evidence that EAL was a risk factor for emergent numeracy or the other 2 outcomes, when the influence of other factors is taken into account.

The study collected information on the frequency of use of digital devices by asking parents how often their child used a desktop or laptop computer, tablet device or smartphone with the response options of 'never or hardly ever', 'at least once a month, but not every week', 'at least once a week, but not every day' and 'every day'. Based on the previous bivariate analysis, Kettlewell and others (2020b) reported that low use of digital devices was associated with greater development in emergent literacy.

The multivariate analysis reported here showed that using a digital device once a month but not more frequently (accounting for 9% of children (Kettlewell and others, 2020b)) was associated with greater development in emergent literacy compared with never or hardly ever using a digital device (6%). The difference in development was equivalent to approximately 4 months, with an effect size of 0.32. Children who used a digital device more frequently (at least once a week and every day) did not have significantly greater or lower development than those who never used a digital device, suggesting that using one more frequently does not have a significant relationship with children's emergent literacy development.

It is important to note that this study was conducted prior to the Covid-19 pandemic. Given the move to remote learning and the periods of country lockdown it is likely that digital device use among young children will have increased, particularly for leisure purposes (see Andrew and others, 2020).

Emergent literacy – no association with gender or back-and-forth conversations

The IELS national report for England (Kettlewell and others, 2020b), which examined the relationship between single variables and the learning outcomes, found that girls were ahead of boys in emergent literacy. However, gender was not identified as a protective factor in the multivariate model for emergent literacy. Whilst it may be significantly related to development on its own, when gender was entered into the model for emergent literacy, it did not add any predictive power over and above the variables already in the model and therefore it was not included in the final model. This suggests that there are

other characteristics which explain more about the differences in children's emergent literacy development than gender.

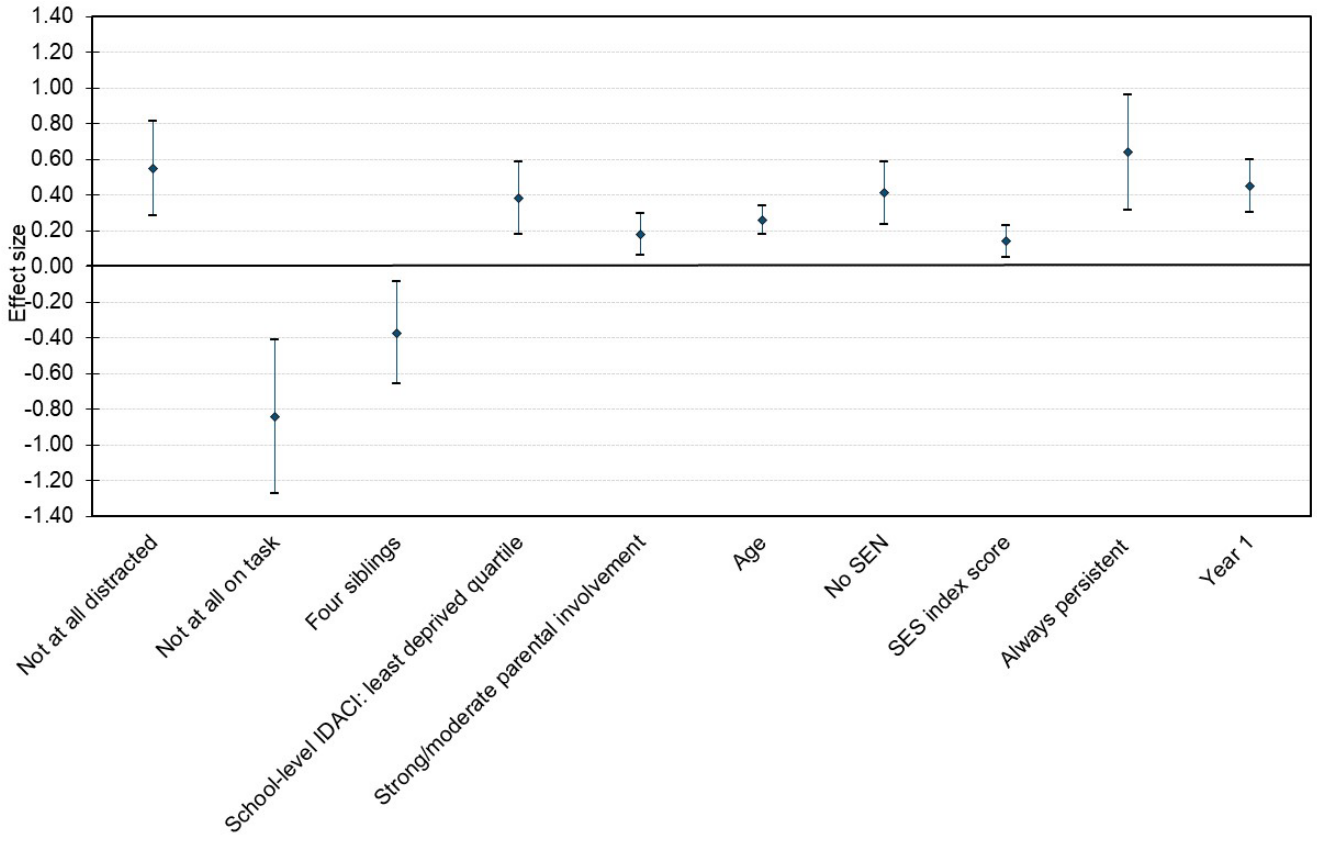
IELS gathered information on children's home learning environment (HLE) through a set of questions in the parent questionnaire. The OECD report for England (OECD, 2019b) found that, after controlling for SES, children whose parents had a back-and-forth conversation with them about their feelings had greater development in emergent literacy and emotion identification. However, having regular back-and-forth conversations about how children were feeling was not identified as a protective factor in the multivariate model for emergent literacy conducted in this analysis. This suggests that there are other characteristics which explain more about the differences in children's emergent literacy development than having back-and-forth conversations with children about their feelings.

Other factors, such as SEN and parental involvement are discussed later in this chapter.

Factors associated with emergent numeracy development

Figure 6 shows the factors that were significantly associated with children's development in emergent numeracy. As for emergent literacy, the factors with the strongest relationship with emergent numeracy (indicated by the effect size) were on-task behaviour, distraction and persistence.

Figure 6: Factors associated with emergent numeracy development



Source: IELS assessment of 2,577 children, age 5

Emergent numeracy – on task, distractibility and persistence

Children who were rated as being on task to a large extent during the direct assessments had greater development in emergent numeracy compared with the small minority of children who were not at all on task. The effect size was -0.84, which is equivalent to approximately 9 months' difference in emergent numeracy. The effect of being not at all on task was associated with larger confidence intervals, the upper confidence interval was -0.41 and the lower confidence interval was -1.27. There were also significant associations between greater development in emergent numeracy for the children rated as on task to a large extent compared to those who were not really on task and also compared to those who were on task to some degree (see Appendix C for further details of these relationships.

Children who were rated by study administrators as being not at all distracted during the direct assessments had significantly greater development in emergent numeracy than the children who were rated as being distracted to a large extent. The effect size for this relationship was 0.55, which is equivalent to approximately 7 months.

The difference between the children who were always persistent compared with those who were never persistent had an effect size of 0.64, which is equivalent to around 7 months. There was also a significant difference in children's emergent numeracy development between the children rated as often persistent compared with those who were never persistent, where children rated as often persistence showed greater development – see Appendix C for further details.

As for emergent literacy, deprivation measures of school-level IDACI and parental SES had relatively lower effect sizes for emergent numeracy than distraction, on-task behaviour, persistence, having an identified special educational need (SEN) and being in Year 1.

Emergent numeracy – having 4 siblings

There was one risk factor that was unique to the model for emergent numeracy, namely having 4 siblings.

The OECD national report for England (OECD, 2020b) reported that development in emergent literacy and emergent numeracy for children with 1 sibling did not differ significantly from those of children with no siblings or those with 2, after accounting for socio-economic status. However, having 4 or more siblings was associated with lower emergent literacy and numeracy development (OECD, 2020b).

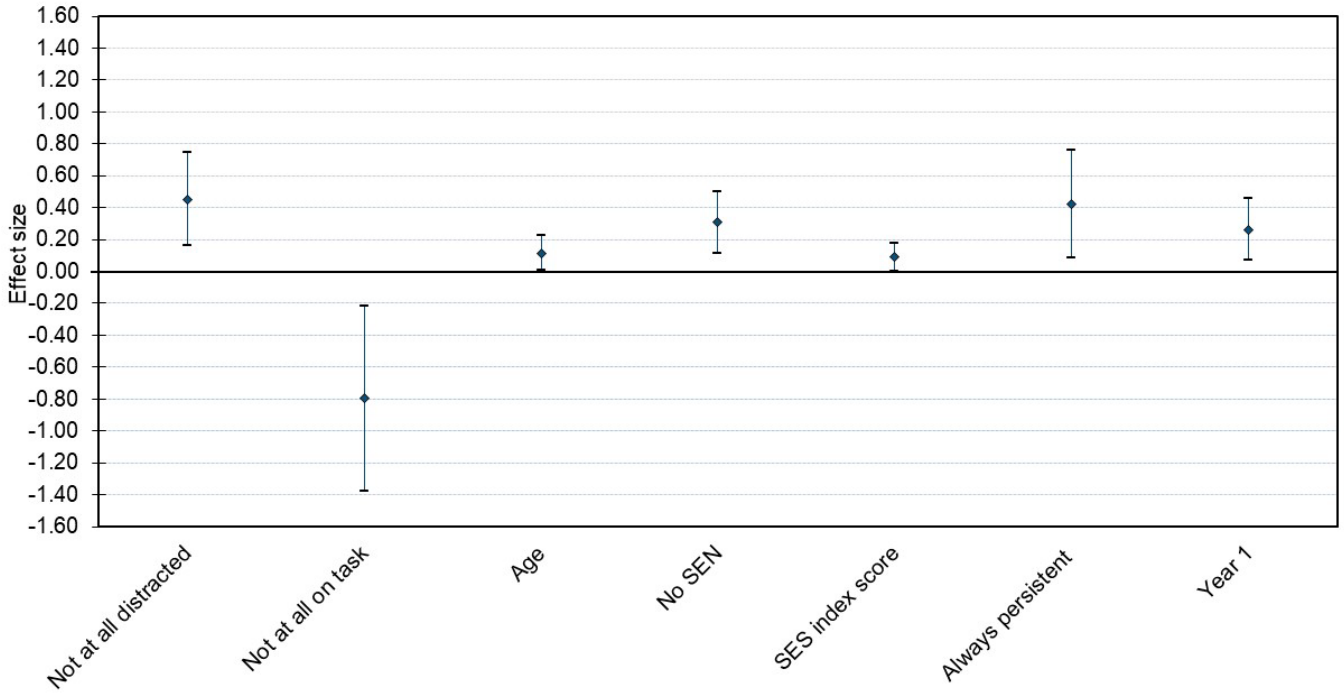
In this analysis, when all other factors were considered, having 4 siblings (but not more) was associated with lower development in emergent numeracy, with an effect size of 0.37. Children who had 4 siblings (3% of children) had a lower level of development, equivalent to 4 months, compared with children who had no siblings (17%). This relationship may be associated with the implications of parents having to split their energy, time and money across more children (Downey, 2001) as family size is a strong predictor of parental investment even for families with a higher socio-economic status (Lawson and Mace, 2009). However, it is worth noting that this analysis did not find a relationship between the number of siblings and development in any other learning outcome nor was it seen for children who had more than 4 siblings.

Other factors, such as SEN and parental involvement are discussed later in this chapter.

Factor associated with mental flexibility development

Figure 7 shows the factors that were significantly associated with children's development in mental flexibility.

Figure 7: Factors associated with mental flexibility development



Source: IELS assessment of 2,577 children, age 5

Mental flexibility – on task, distractibility and persistence

As for emergent literacy and numeracy, the factors with the strongest relationship with mental flexibility (indicated by the effect size) were on-task behaviour, distraction and persistence.

Children who were rated as being on task to a large extent during the direct assessments compared with the small minority of children who were not at all on task was -0.80, which was equivalent to approximately 14 months' difference in mental flexibility. The effect of being not at all on task was, again, associated with larger confidence intervals, the upper confidence interval was -0.22 and the lower confidence interval was -1.38. There were also significant associations between greater development in mental flexibility for the children rated as on task to a large extent compared to those who were not really on task and also compared to those children who were on task to some degree (see Appendix C for further details of these relationships).

Children who were rated by study administrators as being not at all distracted during the direct assessments had significantly greater development in mental flexibility than the children who were rated as being distracted to a large extent. The effect size for this relationship was 0.45, which is equivalent to approximately 8 months' difference in mental flexibility.

The difference between the children who were always persistent compared with those who are never persistent had an effect size of 0.42, which is equivalent to around 7 months. There was also a significant difference in children's mental flexibility development between children rated as often persistent compared with those who were never persistent – see Appendix C for further details.

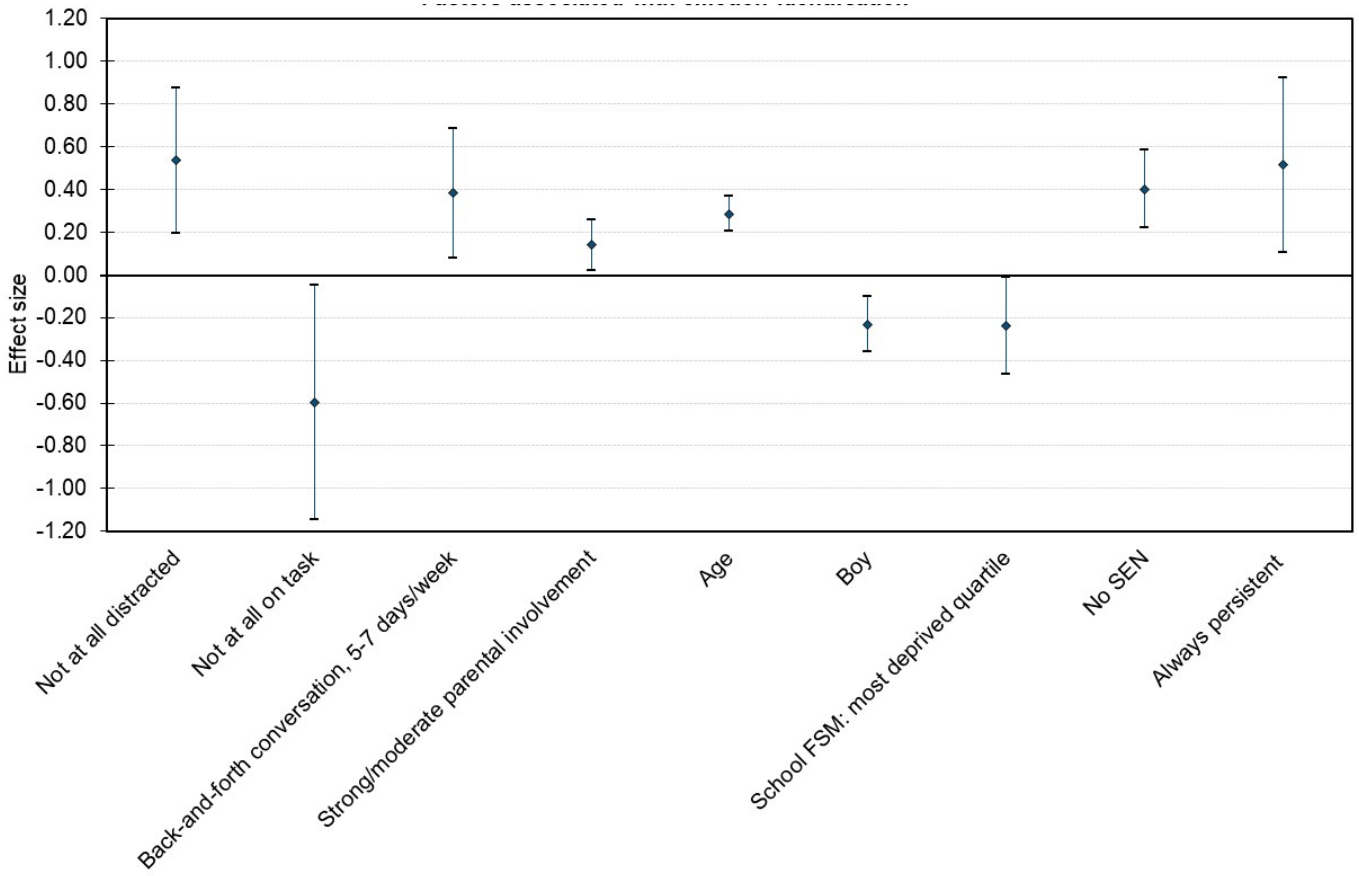
The deprivation measure that was significantly related to mental flexibility (child SES) had the lowest effect size out of all the 7 factors identified in the model. There were no risk or protective factors that were unique to mental flexibility.

Other factors, such as SEN are discussed later in this chapter.

Factors associated with emotion identification development

Figure 8 shows the factors that were significantly associated with children's development in emotion identification.

Figure 8: Factors associated with emotion identification development



Source: IELS assessment of 2,577 children, age 5

Emotion identification – on task, distractibility and persistence

As for the other 3 learning outcomes, the factors with the strongest relationship with emotion identification (indicated by the effect size) were on-task behaviour, distraction and persistence.

Children who were rated as being on task to a large extent during the direct assessments compared with the small minority of children who were not at all on task was -0.60, which was equivalent to approximately 9 months' difference in emotion identification. The effect of being not at all on task was associated with larger confidence intervals, the upper confidence interval was -0.05 and the lower confidence interval was -1.15.

Children who were rated by study administrators as being not at all distracted during the direct assessments had significantly greater development in emotion identification than the children who were rated as being distracted to a large extent. The effect size for this relationship was 0.53, which is equivalent to approximately 8 months' difference in emotion identification.

The difference between the children who were always persistent compared with the children who are never persistent had an effect size of 0.51, which is equivalent to around 8 months' difference in emotion identification. The effect of being always persistent was associated with larger confidence intervals, the upper confidence interval was 0.92 and the lower confidence interval was 0.11.

The deprivation measure that was significantly related to emotion identification (being in the highest quartile for school FSM) had a relatively low effect size compared with the other factors included in the model.

Emotion identification – gender and back-and-forth conversations

Two factors that were significant in the model were unique to emotion identification, namely: gender and having a back-and-forth conversation with parents about the child's feelings.

The 2018 EYFSP results (DfE, 2018a) show evidence of a gender gap in children's social-emotional development at age 4, with more girls than boys reaching the expected level in all 3 ELGs for personal, social and emotional development (PSED). The IELS national report for England (Kettlewell and others, 2020b), which examined the relationship between single variables and the learning outcomes, found that girls were ahead of boys in emotion identification and emergent literacy. In this current analysis, when all other factors are considered, gender was significantly related to emotion identification, with an effect size of 0.23. Girls had greater development in emotion identification compared to boys, equivalent to around 3 months.

IELS gathered information on children's home learning environment (HLE) through a set of questions in the parent questionnaire. The OECD report for England (OECD, 2019b) found that, after controlling for SES, children whose parents had a back-and-forth conversation with them about their feelings had greater development in emergent literacy and emotion identification.

The multivariate analysis reported here found, when all other factors were considered, greater development in children's emotion identification was significantly associated with the frequency of their parents having back-and-forth conversations about children's feelings. Children whose parents had such conversations with them less than once a week (4% of children) had lower development than those who had back-and-forth conversations 5 to 7 days a week (52%), 3 to 4 days a week (29%) and 1 to 2 days a week (15%). The difference in development between children who never have back-and-forth conversations and those who have conversations 5 to 7 days a week had an effect size of 0.38, which is equivalent to around 6 months' difference in emotion identification.

Other factors, such as SEN and parental involvement are discussed below.

Protective factors associated with 5-year-olds' development

The risk and protective factors most commonly associated with children's learning outcomes are discussed in more detail below.

Several protective factors were significantly related to at least 3 of the measures, including being on task/not distracted, persistence, age, year group and parental involvement with activities at their child's schools. A full list of all the factors in the models can be found in Appendix C.

When discussing factors associated with development it is useful to consider those which are fixed, that is, the variable itself cannot be changed; and those which are malleable, that is, they are changeable characteristics (Gorard and Huat See, 2013). If a factor is fixed, it does not mean that policy and intervention are not necessary, but that the characteristic may be best used as an identifier.

Fixed protective factors

Age and year group

Children in the IELS sample in England ranged from 4 years 11 months to 6 years 0 months at the time of assessment. The sample was also split across 2 year groups: Reception and Year 1, with the majority of children in the sample in Year 1 (84% (Kettlewell and others, 2020b). The IELS age variable devised by the OECD was calculated as the difference between the date of birth and the date of assessment. The statistical analysis reported here found that children's age in months was found to be

significantly associated with all 4 learning outcomes. The effect size associated with being older, for example 6 years 0 months old compared to 5 years 0 months old at the time of assessment, and was 0.19 for emergent literacy, 0.26 for emergent numeracy, 0.11 for mental flexibility and 0.29 for emotion identification. This supports the findings in the IELS national report for England (Kettlewell and others, 2020b), which used bivariate analysis, that there was a significant difference in children's development between children who were the oldest (6 years, 0 months) at the time of the study and those who were the youngest (4 years, 11 months).

In the multivariate analysis reported here, being in Year 1 was also significantly associated with greater development in emergent literacy, emergent numeracy and mental flexibility, with the effect sizes 0.37, 0.45 and 0.26 respectively, indicating small effects. The effect of being in Year 1 is true over and above the effect of children being older. This finding is not surprising given that, these children would have had an extra year of schooling by the time of the study and these outcomes might be expected to be enhanced by their additional experience in school.

Malleable protective factors

Persistence and staying on task

Persistence is considered to be an important attribute for success in education. Research has identified a relationship between persistence and later cognitive development (Banerjee and others, 2007; Mokrova and others, 2013). Persistence and a passion for long-term goals are also associated with success in later life (Duckworth and others, 2007).

Persistence was found to be significantly associated with greater development in all 4 learning outcomes. The difference between the 6% of children who are always persistent compared with the 4% who are never persistent was equivalent to around 7 months in emergent literacy, emergent numeracy and mental flexibility, and 8 months in emotion identification). A significant difference in children's emergent numeracy and mental flexibility development was also seen between the 27% of children rated as often persistent compared with the 4% who were never persistent – see Appendix C for further details.

Although persistence may be viewed as an innate personality trait, research evidence suggests that it can be influenced, meaning it is also malleable. Leonard and others (2019) studied task persistence among a group of 4- and 5-year old children. The study found that adults' actions, outcomes and words all had an impact on a child's persistence. Even when children were told the task was hard, as long as an adult demonstrated it was achievable, they would persist. This highlights the importance of adult role models in demonstrating the benefits of persistence to children, both at home and in school.

The effect sizes of being on task to a large extent were greater than any other factor for each learning outcome, and both factors (being on task to a large extent and being not at all distracted) were significantly related to greater development in all 4 learning outcomes. For example, the effect of being on task to a large extent ranged from 0.60 for emotion identification to 0.84 for emergent numeracy. The difference between the 60% of children who were on task to a large extent during the assessment and the small minority (1%) of children who were not at all on task was equivalent to approximately 9 months in emergent numeracy and emotion identification, 10 months' for emergent literacy and 14 months' for mental flexibility. An association was also seen between greater development in emergent literacy, emergent numeracy, mental flexibility and emotion identification for the 60% of children rated as on task to a large extent compared to the 6% who were not really on task. There was even a significant difference between the 60% of children who were on task to a large extent and the 32% who were on task to some degree in development in emergent literacy, emergent numeracy and mental flexibility.

While it may not seem surprising that children who were focused on the tasks in IELS were able to demonstrate greater development compared with children who were not focused, there is evidence that attention is important in its own right. Previous research has found that both persistence and the ability to stay on task are important aspects of a child's ability to learn (EEF, 2018; Vitiello and others, 2011). Attention and persistence can provide external indicators of children's cognitive flexibility (part of the brain's 'executive function', Miyake and others, 2000). Cognitive flexibility and persistence require an element of attention, both immediately and in the longer-term: children who are more cognitively flexible are able to regulate their attention and children who are persistent are better able to maintain their attention over time when working towards a goal. This may help to explain why both children's persistence and their attention during the direct assessments were significantly associated with greater development in all 4 learning outcomes.

Parental involvement in school activities

Teachers were asked to rate how involved each child's parents were in activities taking place at the school, such as school fetes, concerts and parents' evenings. This measure was confined to parents' involvement in school, rather than their wider engagement with their children's learning, though the two are related and it is known that parental engagement is positively related to children's attainment (Gorard and Huat See, 2013; EEF, 2018). The analysis found that the 69% of children whose parents were strongly or moderately involved with their child's school (Kettlewell and others, 2020b) had greater development in emergent literacy, emergent numeracy and emotion identification, equivalent to around 2 months' worth of schooling in each learning outcome (with effect sizes of 0.18, 0.18 and 0.14 respectively).

Hoover-Dempsey and Sandler's model of parental involvement processes (1995, 1997, and 2005) identifies 3 motivations for why parents become involved in their children's education. First is parent's motivational beliefs relevant to involvement, including both parents' beliefs about what they are supposed to be doing (role construction) and a belief in one's own abilities to act in a certain way will lead to a desired outcome (self-efficacy for helping to succeed). Secondly, the model recognises a key motivation is invitation to involvement from others such as schools, teachers, and the child. Finally, parents' life contexts play a role in their level of involvement including their skills and knowledge for involvement, their energy for involvement and their available time. This suggests a role for schools and teachers in supporting parents to feel they are invited to be involved and also to communicate clearly what schools expect of parents.

Harris and Goodall (2007) argue that one of the greatest barriers to parental engagement is parents' own experiences of education, which is likely to be linked to SES. Studies have suggested that increased parental involvement goes some way to off-set the detrimental effects caused by deprivation (Hango, 2005), reduces the likelihood of later school drop-out (Barnard, 2004) and has a positive effect on achievement (Barnard, 2004; Houtenville and Conway, 2008).

Risk factors associated with 5-year-olds' development

Risk factors that are associated with lower development for all 5-year-olds in our sample for all or most of the 4 learning outcomes are discussed below. A full list of all the factors in the models can be found in Appendix C.

Fixed risk factors – child and family characteristics

Having an identified special educational need (SEN)

As reported in Kettlewell and others (2020b), 12% of the children in the IELS sample had a SEN[28] identified in the national pupil database. The majority of these children (61% of the 299 who had an identified SEN) had difficulties with communication and interaction (including speech, language and communication difficulties and autistic spectrum disorder). In addition, 14% had difficulties with cognition and learning (moderate learning difficulties and specific learning difficulties); 13% had social, mental and emotional health issues; 4% had sensory and/or physical needs (hearing impairment, visual impairment and physical disability); and 8% had other difficulties or no specialist assessment of the type of need.[29]

[28] The SEN category in IELS is representative of SEN children who are enrolled in mainstream schools, therefore it would be expected that children with more severe SEN are not represented in the sample because special schools were not sampled, and in addition because any children with an SEN/disability severe enough to prevent them from engaging with the assessments were not asked to participate.
[29] The number of children in each category was too small to allow for analysis by type of SEN.

The IELS national report for England (Kettlewell and others, 2020b), based on bivariate analysis, reported that children with an identified SEN showed markedly lower scores across all measures with the exception of trust, in which they showed higher levels when compared with children with no identified SEN. Previous evidence from the 2019 EYFSP results (DfE, 2019a) has found pupils with an identified SEN are less likely to have reached at least the expected standard in all ELGs.

In this research using multivariate analysis, having an identified SEN at the age of 5 was identified as a significant risk factor for all 4 learning outcomes, with effect sizes ranging from 0.31 for mental flexibility to 0.46 for emergent literacy. Children identified as having a SEN had lower development equivalent to around 4 months in emergent numeracy, 5 months in mental flexibility and 6 months in emotion identification and emergent literacy.

Having an identified SEN was significantly correlated with deprivation, as measured by parental SES. The relationship between SEN and deprivation is discussed in more detail below.

Amount of child-level variation explained by the final models

The amount of child-level variation explained by the final models provides an additional piece of information which can be used to understand how informative the individual models are, and to allow comparisons between models. (Details about how the percentage of variance explained by each model is calculated can be found in Appendix A). It is rare to have a model that explains over 70% of the variation in a learning outcome and this is normally only achieved with a reliable measure of prior attainment. Models that explain below 50% are quite normal and ultimately helps to conclude that there are other factors, that have not been measured, that explain more of the variation in the outcome of interest.

Table 5 shows how much of the variation in each learning outcome can been explained by the independent variables included in each model. For example, the emergent numeracy model explains 45% of the variation in development which means that the remaining 55% of variation is explained by factors for which the IELS study did not collect information on. There are many other factors which are related to deprivation which have been shown to impact on development, including stress in pregnancy, children's attachment, adverse childhood experiences, inter-parental conflict and the impact of intergenerational deprivation (Crenna-Jennings, 2018) which are not captured by this study (and indeed are difficult to capture in a cross-sectional study of this kind).

Table 5: Percentage of variation in children's learning outcomes explained by each model

Emergent literacy	Emergent numeracy	Mental flexibility	Emotion identification
41%	45%	23%	25%

Source: IELS assessment of 2,577 children, age 5

The models explain more variation for emergent literacy and emergent numeracy than they do for emotion identification and mental flexibility. This was also true for the analysis using IELS data to investigate children's physical development (see Lucas and others, 2021).

Do the risk and protective factors differ by deprivation?

To investigate the relationship between deprivation and children's development further, the same analysis was conducted with children who were in the most and least deprived quartiles (using the child-level IDACI measure). Child-level IDACI was chosen instead of the other deprivation measures as it is an established deprivation measure which is widely used in research conducted in England. As a continuous measure it enables a more detailed analysis than the binary measure of Free School Meals (FSM) and allows an exploration of area deprivation at the child level rather than the school level. Finally, when choosing between child-level IDACI and parental SES, it was decided to use child-level IDACI (which originated from the NPD) instead of the parental SES measure which originated from the parent questionnaire. This was because a child-level IDACI score was available for the majority of children in the sample and therefore required little imputation, whereas 37% of the parental SES scores were imputed.

The child-level IDACI measure was divided into quartiles, resulting in a sample of 644 children living in the most deprived areas and 644 children living in the least deprived areas. This analysis used the same methodology as outlined above, to identify the protective and risk factors of these 2 groups of children and whether they differ.

The amount of variance explained by the models is largely similar to the models including the whole sample of children, as shown in Table 6.

Table 6: Percentage of variation explained by models confined to the most and least deprived groups

IDACI quartile	Emergent literacy	Emergent numeracy	Mental flexibility	Emotion identification
Most deprived	37%	46%	24%	26%
Least deprived	42%	41%	14%	21%

Source: IELS assessment of 1,288 children, age 5

Table 7 shows the risk and protective factors for children living in the most and least deprived areas.

Table 7: Risk and protective factors for children living in the most and least deprived areas

Learning outcome	Factors which were significant for both groups (that is, children living in areas of high and low deprivation)	Factors only associated with development for children living in the most deprived areas	Factors only associated with development for children living in the least deprived areas
Emergent literacy	Being on task Having a special educational need (SEN)	Having 50 books or more at home Being in Year 1 Having English as an additional language (EAL)	Not being distracted Being older Low parental socio-economic status (SES)
Emergent numeracy	Being on task Being older Being in Year 1	Having 50 books or more at home Being persistent Having a special educational need (SEN)	Not being distracted Low parental socio-economic status (SES)
Mental flexibility	Being on task Not being distracted	Being older Having a special educational need (SEN)	-

Emotion identification	Being older Having a special educational need (SEN)	Being on task Being a girl	-

Source: IELS assessment of 1,288 children, age 5

The protective and risk factors associated with development for children in both IDACI quartiles are summarised below, focusing on where there are similarities and differences with the models for the overall sample. Charts presenting the effect sizes for each of the factors and full details on the variables included in each model can be found in Appendix C.

As noted in Chapter 2 where there were age-related differences found in the learning outcome models, this was used to calculate the average gain in points for each additional month of age. Note that for children living in the least deprived areas, age in months was not in the final model for mental flexibility which means the estimate of months' difference cannot be reported.

Protective factors associated with greater development for children living in the most and least deprived areas

This more in-depth analysis provided further support for some of the protective factors identified in the whole sample (namely being older and being on task/not distracted), with limited evidence that these factors affected the most and least deprived children differently. As these protective factors were largely consistent across models, it suggests they are not confined to comparisons between extremes of deprivation.

Being older and in Year 1

For children living in the most and least deprived areas, being older was associated with greater development but not across all the learning outcomes (that is, in all but mental flexibility for those living in the least deprived areas and in all but emergent literacy for children living in the most deprived areas). Similarly, being in Year 1 was associated with greater development in emergent literacy and emergent numeracy for children living in the most deprived areas and just emergent numeracy for children living in the least deprived areas.

Staying on task

For children living in the most deprived areas, being on task to a large extent (compared with not really being on task) was associated with greater development in emergent literacy, emergent numeracy, mental flexibility and emotion identification. This was the strongest protective factor for development in emergent literacy, mental flexibility and emotion identification, indicated by the largest effect size relative to the other factors. For

these children living in areas of high deprivation, greater emergent numeracy development was also associated with being on task to a large extent compared with not being on task at all. This was the strongest protective factor for development in emergent numeracy.

For the children living in the least deprived areas, being on task to a large extent (compared with not being on task at all) was the strongest protective factor in emergent literacy, emergent numeracy and mental flexibility, but not emotion identification.

The analysis revealed that there was a relationship between deprivation and concentration during the direct assessments. A higher proportion of children living in the most deprived areas were rated by the study administrators as not on task. Of the children who were 'not really or not at all on task', 28% were living in the most deprived areas, compared to 21% living in the least deprived areas. Similarly, there was a relationship between deprivation and distraction. Of the children who were rated as distracted 'to a large extent' during the direct assessments, 31% were living in the most deprived areas compared with 19% living in the least deprived areas).

Protective factors specifically for children living in the most deprived areas

Being a girl

Being a girl was a protective factor for children's emotion identification development for children living in the most deprived areas, as was seen in the model for all children. This difference was equivalent to approximately 4 months, with an effect size of 0.24.

There was no significant relationship between gender and emotion identification among children living in the least deprived areas.

Persistence was a protective factor for children living in the most deprived areas

Previous research suggests that children's persistence is related to secure attachment with parents and caregivers in the early months of life (Drake and others, 2013). In addition, there is evidence that children's cognitive function is related to family deprivation. Children from households with a lower SES tend to show lower development in executive function and cognitive self-regulation (McClelland and others, 2015; Hackman and others, 2015; Ackerman and FriedmanKrauss, 2017). When focusing exclusively on children living in the most deprived areas, persistence appeared to be important for their emergent numeracy development but was not significantly associated with greater development in the 3 other learning outcomes (though analysis of all children in the sample found that persistence was a protective factor for all 4 learning outcomes). Children living in the most deprived areas who were rated as always persistent (4%) had greater development in emergent numeracy than those who were never persistent (6%), equivalent to approximately 8 months' difference and with an effect size of 0.76.

This analysis found no significant relationship between persistence and the learning outcomes for children in the least deprived group.

Previous research has found a relationship between an individual's view of their own efficacy and educational outcomes, and persistence is known to be related to an individual's locus of control (Crenna-Jennings, 2018; Flouri, 2006). For older children, believing you are in control of your intelligence and having a stronger sense of agency is seen more among disadvantaged, high-achieving pupils than it is by their low-achieving peers who are also disadvantaged (Siraj-Blatchford and others, 2011; Classick and others, 2021).

Number of books in the home for children living in the most deprived areas

The number of books in the home has often been used as a proxy for socio-economic status in large-scale international assessments such as the Progress in International Reading Literacy Study (Mullis and Martin, 2015). Research suggests that children who are eligible for FSM tend to have fewer children's books in the home (Wheater and others, 2020). The IELS parent questionnaire included questions on children's home learning environment, including the number of children's books in the home.

The analysis reported here found that for children living in the most deprived areas, having more books in the home was significantly associated with greater development in emergent literacy and emergent numeracy. Children in the most deprived child-level IDACI quartile who had more than 100 books in their home (which accounted for 14% of those living in the most deprived areas) had significantly greater development in emergent numeracy and emergent literacy than those who had 10 children's books or fewer in the home (16% of those living in the most deprived areas). The effect sizes were 0.58 for emergent literacy and 0.53 for emergent numeracy and the difference was equivalent to approximately 5 months in emergent numeracy and around 8 months in emergent literacy. A similar positive relationship was also found for children who had 51 to 100 books (23% of children living in the most deprived areas). The availability of books in the home is another way the IELS study measured cultural capital. The availability of books in the home, book-sharing activities, and high-quality linguistic interactions, between child and caregiver are associated with improved language outcomes such as vocabulary and early literacy development (Asmussen and others, 2018). This analysis suggests that access to children's books may be particularly related to emergent literacy for children living in the most deprived areas.

Risk factors associated with lower development for children living in the most and least deprived areas

As in the main sample, the analysis revealed fewer risk than protective factors. The following section discusses the similarities and differences in risk factors between the models for all children and those living in the most and least deprived areas.

Special educational needs (SEN)

For children living in the most deprived areas, having an identified SEN was associated with lower levels of development in emergent literacy, emergent numeracy, mental flexibility and emotion identification. For these children, the difference between children with an identified SEN and children without is equivalent to around 4 months in emergent numeracy, 6 months for emergent literacy and mental flexibility and 7 months for emotion identification. For children living in the least deprived areas, having an identified SEN was a significant risk factor for emergent literacy and emotion identification only, with differences of approximately 6 and 7 months respectively. This is unexpected, as SEN was a significant risk factor for all 4 outcomes in the model based on all children. However, this finding may have been influenced by the relationship between SEN and deprivation in the IELS sample. Overall, 12% of children in the IELS sample in England had an identified SEN according to the NPD (Kettlewell and others, 2020b). The incidence of SEN was lower among children from the least deprived families (8% of children living in the least deprived areas had an identified SEN, compared to 14% of children living in the most deprived areas).

Risk factors specifically for children living in the most deprived areas

English as an additional language

As in the previous analysis of the whole sample, EAL was associated with lower levels of emergent literacy among children living in the most deprived areas, but was not a significant risk factor for children living in the least deprived areas. This may be influenced by the fact that having English as an additional language was more common in the areas of high deprivation than the areas of low deprivation. A larger proportion of children (24%) in the most deprived IDACI quartile had EAL compared with 8% of children in the least deprived IDACI quartile.

Risk factors specifically for children living in the least deprived areas

Relationship with parental SES for children living in the least deprived areas

For children living in the least deprived areas, another deprivation measure (that is, low parental SES) was a risk factor associated with lower development in emergent literacy and emergent numeracy, although the effect size was small at 0.19 for both learning

outcomes. For the children living in the most deprived areas, an association with parental SES was not seen in the final models.

While 41% of children living in the least deprived areas were in the least deprived parental SES quartile, 7% of children who were living in the least deprived areas were in the most deprived SES quartile, according to data from the parent questionnaire. This highlights the differences between the two measures, as discussed in Chapter 2. The parental SES measure is based on household income, parental occupation and parental education whereas child-level IDACI is based upon the proportion of the population in the area the child lives who are experiencing deprivation related to low income.

Based on this analysis alone is it not clear what is driving this association. As noted above, it could be related to the differences in the measures. Alternatively, it could be that the differences in emergent literacy and emergent numeracy development between the children living in the least deprived areas is driven by comparatively lower development for low parental SES children or comparatively greater development for children with higher parental SES. Either way, this finding warrants further investigation.

Which variables would be expected to relate to development but were not found in this analysis?

There were some variables which did not have a significant relationship with the learning outcomes, despite expectations to the contrary. This was particularly the case for some features of the HLE and low birthweight.

The home learning environment

The HLE is an important factor in the development of early speech, language and communication (Sylva and others, 2004; DfE and the National Literacy Trust, 2018b). Children's early numeracy and number development is also consistently associated with the quality of the HLE (Sylva and others, 2004; Asmussen and others, 2018). A higher quality HLE is more prevalent in families with higher socio-economic status (Law and others, 2017; Sim and others, 2018). Deprived children are less likely to experience a HLE that supports their early cognitive development, particularly in early literacy and language (DfE, 2017). However, a high-quality HLE also operates independently from social class, which means that children from deprived backgrounds with a high quality HLE have better outcomes than children from deprived backgrounds with a lower quality HLE (Melhuish, 2010; Sylva and others, 2007).

Previous reports on the IELS findings by Kettlewell and others (2020b) and OECD (2020b) found positive relationships between several features of the HLE and children's development including helping children to read, having back-and-forth conversations

about their children's feelings, attending special or paid for activities and the number of children's books in the home.

The further analysis reported here identified relationships between children's outcomes and certain HLE variables (such as having back-and-forth conversations, the number of children's books in the home and the children's use of digital devices at home), but not others such as reading to children and helping them to read. It is possible that these results reflect an equalising effect of schooling on the HLE. For example, most primary schools set expectations for parents to read to their children and help them to read. If this is the case, then there would be more uniformity in children's HLE in Reception and Year 1 than before they start school, and this could be why its influence on children's development is not apparent in the models, even though it is strongly supported in previous research (Melhuish and Gardiner, 2018; Melhuish and Gardiner, 2020; Sylva and others 2008). On the other hand, the analysis found that parental involvement in school activities was significantly related to learning outcomes. It is possible that parents who are more involved in their child's school also provide a stronger HLE than those who are slightly or not involved, in which case the HLE variables that were not identified as significant in the models did not add any more information above parental involvement with school and the other variables included in the models.

Low birthweight

The previous report, using bivariate analysis (Kettlewell and others, 2020b) found that the 11% of the children in the IELS sample with a low birthweight (2.5kg[30] or less) had significantly lower outcomes in emergent literacy and emergent numeracy than their counterparts who did not have low birthweight.

However, having a low birthweight did not feature as a risk factor in the current analysis, once the influence of other variables had been taken into account. This does not appear to be due to a relationship between deprivation and low birthweight because children with low birthweight were roughly equally distributed across the child-level IDACI quartiles, and the OECD analysis found no relationship between the IELS measure of parental SES and low birthweight in England (OECD, 2020b). However, it is worth noting that low birthweight was particularly hard to impute with accuracy, in comparison to the other variables from the parent questionnaire. It could be that this is driving the absence of low birthweight as a risk factor, but further investigation would be needed to establish this with certainty.

[30] Equivalent to 5 pounds 8 ounces.

Early childhood education and care (ECEC)

In England, some 2-year-olds, and all 3-, and 4-year olds, are entitled to free part-time early childhood education and care (ECEC). At the age of 2, children from deprived backgrounds are eligible for 15 hours of free ECEC per week[31]. Existing research suggests that ECEC provision can play an important role in supporting children's early cognitive development (Sylva and others, 2008; Bonetti and Brown, 2018; Melhuish and Gardiner, 2020). Research suggests that deprived children have the most to gain from ECEC. While deprived children are less likely to be in ECEC, particularly centre-based care, they were found to benefit significantly from good quality ECEC experiences, particular when they are able to mix with children from a range of different social backgrounds (Sylva and others, 2004; Roberts and others, 2010, Albakri and others, 2018).

The parent questionnaire collected extensive information on early childhood education and care (ECEC), including age of attendance, type of setting and intensity of attendance. The IELS study (OECD, 2020b) found few significant differences by ECEC factors after adjusting for parental SES and in the multivariate analysis reported here, ECEC did not feature as a significant protective factor in any of the models. This may reflect the fact that the majority of children in England attended ECEC (98% of children in the IELS sample attended some form of ECEC with 71% of these children first attending before the age of 3 and 29% attending at age 3 or 4 (Kettlewell and others, 2020b)), thereby reducing the variance between children. In addition, the study did not provide information on the quality of the ECEC provision children had experienced, and ECEC quality is known to be an influential factor for children's early development (Sylva and others, 2020; Shuey & Kankaras 2018).

[31] This applies for 38 weeks per year.

4 Discussion and conclusion

Discussion

This report set out to investigate how different measures of socio-economic deprivation interact with each other to predict early learning outcomes and to identify the risk and protective factors for children's learning outcomes at the age of 5. Finally, it sought to identify the factors most predictive of development for children from more deprived areas.

How do different measures of deprivation interact to predict early learning outcomes at age 5?

The deprivation measures of parental SES, school-level IDACI and school-level FSM are associated with children's learning outcomes

Both family and school deprivation were associated with children's learning outcomes. The analysis found that 3 of the 5 measures of socio-economic deprivation investigated in this report (namely parental SES, school-level IDACI and school-level FSM) were significantly associated with one or more early learning outcomes when the influence of other demographic variables was taken into account. Child-level IDACI and child-level FSM did not add any further predictive power to the models once an alternative measure of deprivation was already included.

Parental SES was related to 3 outcomes (emergent literacy, emergent numeracy and mental flexibility); the proportion of income-deprived families in the area of the child's school (school-level IDACI) was related to 2 outcomes (emergent literacy and emergent numeracy); and the proportion of FSM pupils in a school was related to one (emotion identification). Of the 5 deprivation measures analysed, the SES measure combining parents' income, education and employment was the most consistently related to the learning outcomes investigated. This supports the conclusion that measures of deprivation which include parental education and employment alongside income are more beneficial than measures of parental income alone (Sutherland and others, 2015; Ilie and others, 2017; Taylor, 2018). This may be because parental education represents 'cultural capital' (Bourdieu and Passeron, 1990) which parents can pass on to their children.

For the 2 cognitive learning outcomes of emergent literacy and emergent numeracy, both school-level IDACI and parental SES were related to children's development. The effect size of school-level IDACI was relatively larger than for parental SES, which indicates that whilst the influence of parental SES is important, the deprivation of the area around the school has a stronger association with children's cognitive development than parental SES, possibly due to the quality of teaching and/or interaction with the child's peer group (Sylva and others, 2008).

Finally, whilst there is already a wealth of research into deprivation and literacy and numeracy development, somewhat less is known about how deprivation relates to other aspects of children's development. This analysis found that children whose parents had a higher socio-economic status had greater development in mental flexibility. Further research would be needed to investigate how families with a higher SES support the development of children's mental flexibility and to identify effective interventions to support the development of mental flexibility in deprived children.

The deprivation measures of child-level FSM and child-level IDACI had a weaker relationship with children's learning outcomes than parental SES and school-level IDACI

As already noted, parental SES was a more powerful predictor of the relationship between deprivation and children's outcomes than the other 2 child-level variables. The binary nature of FSM, in comparison to the continuous measure of parental SES, makes it less sensitive to differences in the learning outcomes which are related to deprivation (Taylor, 2018). The reason that school-level IDACI had greater predictive power than school-level FSM for emergent literacy and emergent numeracy is likely to be due to the greater ability of IDACI quartiles to reflect the extent of deprivation. However, it is worth noting that school FSM, rather than school-level IDACI was significantly related to emotion identification. This may be due to the ability of school FSM to more accurately reflect the deprivation of children's peers, rather than the deprivation of the local area surrounding the school. This reflects the finding from the longitudinal Effective Pre-school and Primary Education Project that the characteristics of a school's pupil intake can influence children's social outcomes (Sylva and others, 2008).

However, deprivation measures did not have the strongest effect on children's learning outcomes

Being on task and being persistent were both more strongly related to children's development than deprivation in all 4 learning outcomes. This suggest that children's underpinning cognitive/emotional skills, such as being on task and being persistent, are more important than deprivation.

For mental flexibility, the deprivation measure of parental SES had a smaller effect size than any of the other risk or protective factors in the model. For emotion identification, the effect size of school-level FSM was smaller than most other factors including having an identified SEN, being persistent, being older/in Year 1 and being on task/not distracted.

This indicates that children's learning outcomes are influenced by a wide range of factors, of which deprivation is just one.

Area-level measures of deprivation may be more readily available than parental SES for research but do not take account of individual circumstances

Measures of neighbourhood poverty (IDACI) are more readily available for use in future research than parental SES because the measures are created using national indicators. Arguably, the deprivation of the school reflects the deprivation of the local area which may be related to both the quality of the provision and the degree of challenge (both positive and negative) posed by other children in the school. However, some researchers advise caution in using solely area-level measures to determine deprivation because area-level measures can mask differences in individual circumstances (Gorard and Huat See, 2013) and take no account of differences between areas in other respects, such as access to social services, schools and other education opportunities (Ilie and others, 2017).

What are the protective and risk factors associated with early learning outcomes at age 5?

Turning to the second research question, this investigation has identified a set of protective and risk factors for children's development.

Being on task/not distracted, being more persistent and being older were associated with greater development in all 4 learning outcomes

For all children, being on task/not distracted, being more persistent and being older were related to greater development in emergent literacy, emergent numeracy, mental flexibility and emotion identification. The relationship between attention, persistence and executive function is discussed in more detail below.

Some factors were significantly associated with just one particular learning outcome. For example, parents having regular back-and-forth conversations with their children about their feelings was positively associated with emotion identification development, as was being a girl. Children with EAL had lower development in emergent literacy than those who had English as their first language.

Having an identified special educational need (SEN) was consistently identified as a risk factor for children's development

Having an identified SEN at the age of 5 was identified as a significant risk factor for all children across all 4 learning outcomes, with a larger effect than for deprivation. This is consistent with existing evidence, for example, the 2019 EYFSP results (DfE, 2019a) found pupils with an identified SEN are less likely to have reached at least the expected standard in all ELGs.

This analysis demonstrates the importance of identifying children with a SEN as early as possible and supporting their development. This is particularly important given the likely

impact of the Covid-19 lockdowns on children with SEN (Pascal and others, 2020). In research by Ofsted (2020), early leaders reported that children with SEN had not received the support they needed from other professionals due to service closures or managing visits and some children with SEN had struggled to engage with routines at home in the same way they would in an early years setting.

Are these protective and risk factors different for those children living in deprived areas?

Most of the risk and protective factors were the same for children living in the most deprived areas as for all children

Being older and being on task during the direct assessments were all associated with greater development in learning outcomes for children living in the most deprived areas. Persistence was also found to be related to children's emergent numeracy development for those living in deprived areas.

Having an identified SEN was a risk factor for children living in deprived areas across all learning outcomes and for children living in non-deprived areas for emergent literacy and emergent numeracy.

As with all children, there were some factors which were significantly associated with just one particular learning outcome. For children living in the most deprived areas, having EAL was identified as a risk factor for emergent literacy development and being a girl was a protective factor for emotion identification development.

Having more books in the home was associated with greater emergent literacy and numeracy development for children living in deprived areas only

Having more than 50 children's books in the home was protective of greater development in emergent literacy and emergent numeracy compared with having no books in the home for the children living in the most deprived areas. This relationship was not seen in the models for all children nor for those living in the least deprived areas. This suggests that having access to books may be particularly beneficial for children who are the most deprived. The availability of books in the home, book-sharing activities, and high-quality linguistic interactions between child and caregiver are associated with improved language outcomes such as vocabulary and early literacy development (Asmussen and others, 2018).

Some factors have the potential to be influenced by policy and practice

It is important to distinguish between factors that are more malleable and therefore have the potential to be influenced by policy and practice, and those that are more fixed but may be useful as indicators (Gorard and Huat See, 2013). Given that data for the IELS

study was collected from participants on a single occasion (that is, the study used a cross-sectional design), it is not possible to determine whether findings from IELS represent malleable or fixed factors, but it is possible to draw on evidence from other research to understand more about this. A few variables stand out as important for children's development and are potentially open to influence, namely: children's ability to stay on task and resist distractions; their persistence; and parental involvement at their child's school.

Being on task and persistence are related to a child's executive function

Children's ability to stay on task, and not be distracted during the direct assessments and their persistence were all positively related to their development in the 4 learning outcomes examined in this study (these characteristics were also found to be associated with children's physical development – see Lucas and others, 2021). In addition, persistence was associated with greater emergent numeracy development for children living in the most deprived areas.

Children's persistence and parental involvement in school may be considered malleable factors that can be promoted in early years and school settings. For example, Leonard and others (2019) found that adults are able to influence children's persistence with their actions, outcomes, and words. Persistence also appears related to executive function, which evidence suggests can be influenced through classroom-based pedagogy and specific interventions (Diamond, 2012; Serpel and Esposito, 2016; Ackerman and Friedman-Kraus, 2017; Schmidt and others, 2017; EEF, 2021). These findings suggest that schools should adopt practices that are proven to increase children's attention, persistence and on-task behaviour. It is also important for teachers to identify children who struggle with attention and mental flexibility because these issues may be responsive to intervention. They may also be indicators of undiagnosed conditions such as Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder.

Persistence is related to an individual's view of their own efficacy. An internal locus of control, that is the extent to which individuals feel they have control over their own success or failures, is related to educational outcomes (Crenna-Jennings, 2018; Flouri, 2006). Individuals who believe they can exercise control over their futures are more likely to persevere and take action to improve their skills to become successful in future (Hong and others, 1999). When comparing individuals with equivalent ability or intelligence, those who possess determination and persistence are more likely to be successful (Duckworth and others, 2007).

Parental involvement is related to all children's development

Having parents who were strongly or moderately involved with their child's school (as indicated by their participation in school activities such as fetes, concerts and parents' evenings) was a protective factor for 3 outcomes (emergent literacy, emergent numeracy

and emotion identification). It is not possible to identify the direction of this relationship (that is, whether parental involvement leads to better child outcomes) but there is some evidence that intervening in particular ways to improve parental involvement in school and engagement with their children's learning can be effective in raising attainment (Hutchinson and others, 2020; EEF, 2018). Effective forms of school-to-home and home-to-school communication is important in building strong family partnerships (Epstein and others, 1997). Goodall and others (2011) point out that effective parental engagement requires a whole-school and outward-facing approach; focussed training and coaching for staff; clear, specific and targeted information for parents and use of ICT. Specific challenges include involving parents with low level literacy and numeracy skills, English language capability, confidence in dealing with the school or their child's learning and those who choose not to engage.

Harris and Goodall's (2007) finding that parental engagement in children's learning at home makes the greatest difference to educational achievement has particular relevance for the Covid-19 generation. This suggests that schools should continue to involve parents in school life on the basis that parental involvement is a promising avenue for improving children's outcomes (Houtenville and Conway, 2008).

Having English as an additional language is a risk factor for children's emergent literacy development, especially for children living in the most deprived areas

The models for all children and those living in the most deprived areas showed that EAL was a risk factor for emergent literacy development. However, it was not an identified risk factor for children living in the least deprived areas, suggesting children with EAL in less deprived areas do not appear to be disadvantaged by their additional language at home.

Previous research has found that having EAL is associated with lower levels of literacy and numeracy in the early years, but this effect reduces markedly with age and is largely eliminated by age 16 (Strand and others, 2016; Gorard and Siddiqui, 2019; DfE, 2019b). This analysis suggests that targeted support for children with EAL who also live in deprived areas could be particularly beneficial in reducing the literacy gap at an earlier age.

Conclusion

This research sought to understand how deprivation relates to children's development at age 5, and which factors (such as children's individual and family characteristics) are positively or negatively related to children's outcomes, over and above the influence of deprivation.

The study found that measures of deprivation are related to lower development across emergent literacy, emergent numeracy, mental flexibility and emotion identification. School-level IDACI and parental SES index appear to be most useful measures to

capture the relationship between deprivation and early cognitive outcomes. Child-level FSM was not associated with children's outcomes once other variables were included in the models, although school-level FSM was related to emotion identification (an important aspect of empathy), not school-level IDACI. This study indicates that measures of parental job role and educational level add further explanatory power to purely income-related measures of individual deprivation. In relation to the area-based measures, school-level IDACI is more strongly associated with young children's cognitive development than school-level FSM. These insights have the potential to add value to future investigations of the link between deprivation, family demographics and young children's cognitive development.

For all children, being on task/not distracted and being persistent are associated with greater development, whereas having an identified SEN is consistently related to lower development in all 4 measures of emergent literacy, emergent numeracy, mental flexibility and emotion identification.

This analysis is timely, as the national lockdowns due to the Covid-19 pandemic are likely to have disproportionally affected the most vulnerable children, including those living in poverty, SEN and EAL (Pascal and others, 2020; Ofsted 2020). This research suggests that targeting interventions on schools serving the most deprived areas are most likely to reach most of the children in greatest need.

Persistence and the ability to stay on task may be related to a child's cognitive flexibility (part of the brain's executive function). As these characteristics were associated with development across the models and had relatively large effect sizes, further investigation is merited into their relationship with executive function and the effectiveness of interventions designed to improve persistence and attention. Not only could this benefit children's cognitive development, but our analysis also suggests these factors play a role in mental flexibility and emotion identification.

An increased awareness and understanding of the risk and protective factors for all children is beneficial for families, ECEC practitioners, teachers and policy makers. Particular focus should be given to the multiple factors affecting the development of children from deprived backgrounds, such as encouraging persistence and on-task behaviour, and the availability of children's books in the home as well as targeted interventions to support the language development of children with EAL who live in deprived areas.

References

Ackerman, D.J. and Friedman-Krauss, A.H., 2017. Preschoolers' executive function: importance, contributors, research needs and assessment options. *ETS Research Report Series*, [e-journal] 2017(1), pp.1– 24. 10.1002/ets2.12148 .

Albakri, M., Basi, T., Davies, M., Forsyth, E., Hopwood, V., Patel, R., Skipp, A. and Tanner, E., 2018. *Take-up of free early education entitlements.* [pdf] Available at: <take-up of free early education entitlements> [Accessed 10 March 2021].

Allison, P., 2012. *When can you safely ignore multicollinearity?* [online] Available at: <when can you safely ignore multicollinearity?> [Accessed 13 April 2021].

Andrew, A., Cattan, S., Costa-Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A. and Sevilla, A., 2020. *Learning during the lockdown: real-time data on children's experiences during home learning.* [pdf] Available at: <learning during the lockdown: real-time data on children's experiences during home learning> [Accessed 10 March 2021].

Asmussen, K., Law, J., Charlton, J., Acquah, D., Brims, L., Pote, I. and McBride, T., 2018. *Key competencies in early cognitive development: things, people, numbers and words.* [pdf] Available at: <key competencies in early cognitive development: things, people, numbers and words> [Accessed 10 March 2021].

Banerjee, P. and Tamis-LeMonda, C., 2007. Infants' persistence and mothers' teaching as predictors of toddlers' cognitive development. *Infant Behaviour and Development,* [e-journal] 30(2), pp. 479-491. 10.1016/j.infbeh.2006.12.001

Barnard, W., 2004. Parent involvement in elementary school and educational attainment. *Children and Youth Services Review*, [e-journal] 26(1), pp. 39 – 62. 10.1016/j.childyouth.2003.11.002

Bhat, A., Landa R. and Galloway J., 2011. Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Physical Therapy,* [e-journal] 91(7), pp.1116–1129. [pdf] Available at: <current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders [Accessed 25 February 2021].

Bonetti, S. and Brown, K. 2018. *Structural elements of quality early years provision: a review of the evidence.* [pdf] Available at <structural elements of quality early years provision: a review of the evidence> [Accessed 10 March 2021].

Bourdieu, P. and Passeron, J., 1990. *Reproduction in education, society and culture*. London: Sage Publications, Inc.

Charlton, C., Rasbash, J., Browne, W.J., Healy, M. and Cameron, B., 2019. *MLwiN Version 3.04*. Centre for Multilevel Modelling, University of Bristol.

Children's Commissioner, 2020. *Tackling the disadvantage gap during the Covid-19 Crisis.* [pdf] Available at: <tackling the disadvantage gap during the Covid-19 crisis> [Accessed 10 March 2021].

Chowdry, H. and McBride, T., 2017. *Disadvantage, behaviour and cognitive outcomes.* [pdf] Available at: <disadvantage, behaviour and cognitive outcomes> [Accessed 25 February 2021].

Classick, R., Gambhir, G., Liht, J., Sharp, C. and Wheater, R., 2021. *PISA 2018 additional analyses: What differentiates disadvantaged pupils who do well in PISA from those who do not?* [pdf] Available at: <PISA 2018 additional analyses: what differentiates disadvantaged pupils who do well in PISA from those who do not?> [Accessed 18 February 2021].

Crawford, C. and Greaves, E., 2013. *A comparison of commonly used socio-economic indicators: their relationship to educational disadvantage and relevance to Teach First.* [pdf] Available at :< a comparison of commonly used socio-economic indicators: their relationship to educational disadvantage and relevance to Teach First> [Accessed 18 February 2021].

Crenna-Jennings, W., 2018. *Key drivers of the disadvantage gap literature review*. [pdf] Available at <key drivers of the disadvantage gap literature review> [Accessed 10 March 2021].

Department for Education, 2017. *Unlocking talent, fulling potential: a plan for improving social mobility through education.* [pdf] Available at: <unlocking talent, fulling potential: a plan for improving social mobility through education> [Accessed 10 March 2021].

Department for Education and National Literacy Trust, 2018b. *Improving the home learning environment: a behaviour change approach.* [pdf] Available at: <improving the home learning environment: a behaviour change approach> [Accessed 10 March 2021].

Department for Education 2019a. *Early years foundation stage profile (EYFSP) pupil characteristics 2019: Table 1.* [pdf] Available at: < early years foundation stage profile (EYFSP) 2018 additional tables by pupil characteristics> [Accessed 15 February 2021].

Department for Education 2019b. *Key stage 4 performance, 2019 (revised).* [pdf] Available at: <Key stage 4 performance, 2019 (revised)> [Accessed 15 February 2021].

Department for Education 2019c. *Schools, pupils and their characteristics: January 2019.* [pdf] Available at: <Report link to schools, pupils and their characteristics: January 2019> [Accessed 15 February 2021].

Diamond, A., 2012. Activities and programs that improve children's executive functions. *Current Directions in Psychological Science*, [e-journal] 10.1177/0963721412453722

Donders, A., Rogier, T., van der Heijden, G. T., Stijnen, T. and Moons, K. G., 2006. A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, [e-journal] 59(10), pp.1087–1091. 10.1016/j.jclinepi.2006.01.014

Downey, D., 2001. Number of siblings and intellectual development: the resource dilution explanation. *American Psychologies,* [e-journal] 56(7-7), pp.497-504. 10.1080/1081602X.2018.1532310

Drake, K., Belsky, J. and Fearon, R., 2013. From early attachment to engagement with learning in school: the role of self-regulation and persistence. *Developmental Psychology*, [e-journal] 50(5), pp.1350–1361 DOI: 10.1037/a0032779.

Duckworth, A., Peterson, C., Matthews, M. and Kelly, D., 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology,* [e-journal] 92(6), pp.1087-101. DOI: 10.1037/0022-3514.92.6.1087

Education Endowment Foundation, 2018. *Early years toolkit.* [pdf] Available at: <early years toolkit [Accessed 11 March 2021].

Education Endowment Foundation, 2021. Self-regulation strategies. [online] Available at <self-regulation approaches> [Accessed 1 March 2021].

Education Policy Institute 2020. *Preventing the disadvantage gap from increasing during and after the Covid-19 pandemic: proposals from the Education Policy Institute.* [pdf] Available: preventing the disadvantage gap from increasing during and after the Covid-19 pandemic: proposals from the Education Policy Institute [Accessed 18 February, 2021].

Education Select Committee. 2016. *Quality early years education*. [online] Available at: quality early years education [Accessed 10 March 2021].

Epstein, J.L., Coates, L., Salinas, K., Sanders, M.G. and Simon, B.S., 1997. *School, family, and community partnerships: your handbook for action*. Thousand Oaks, CA:Corwin Press.

Feinstein, L., ed., 2015. *Social and emotional learning: skills for life and work.* [pdf] Available at: <social and emotional learning: skills for life and work> [Accessed 15 February 2021].

Flouri, E., 2006. Parental interest in children's education, children's self-esteem and locus of control, and later educational attainment: twenty-six year follow-up of the 1970 British birth cohort. *British Journal of Educational Psychology,* [e-journal] 76(1), pp.41 – 55. 10.1348/000709905X52508

Goodall, J. and Vorhaus, J., 2011. *Review of best practice in parental engagement.* [pdf] Available at: <review of best practice in parental engagement> [Accessed 27 April 2021].

Gorard, S. and Huat See, B., eds., 2013. *Overcoming disadvantage in education.* Oxford: Routledge. 10.4324/9780203741665

Gorard, S. and Siddiqui, N., 2019. *How trajectories of disadvantage help explain school attainment.* 10.1177/2158244018825171

Hackman, D. A., Gallop, R., Evans, G. W. and Farah, M. J., 2015. Socioeconomic status and executive function: developmental trajectories and mediation. *Developmental Science,* [e-journal] 18(5), pp.686-702. [pdf] Available at: <socioeconomic status and executive function: developmental trajectories and mediation> [Accessed 30 April 2021].

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. and Black, W. C., 1995. *Multivariate data analysis*. 3rd ed. New York, NY: Macmillan.

Hango, D., 2005. *Parental investment in childhood and later adult well-being: Can more involved parents offset the effects of socioeconomic disadvantage?* [pdf] Available at: <parental investment in childhood and later adult well-being: can more involved parents offset the effects of socioeconomic disadvantage?> [Accessed 18 February 2021].

Hapkau, C., Isphording, I., Machin, S. and Ruiz-Valenzuela, J., 2020. *Labour market shocks during the Covid-19 pandemic: inequalities and child outcomes.* [pdf] Available at: <labour market shocks during the Covid-19 pandemic: inequalities and child outcomes> [Accessed 18 February 2021].

Harris, A. and Goodall, J., 2007. *Engaging parents in raising achievement do parents know they matter?* Technical report DCSF-RW004. [pdf] Available at:<engaging parents in raising achievement do parents know they matter?> [Accessed 27 April 2021].

HM Government, n.d. *Census geography: output areas and super output areas.* [pdf] Available at: <census geography: output areas and super output areas> [Accessed 10 March 2021].

Hobbs, G. and Vignoles, A., 2009. Is children's free school meal 'eligibility' a good proxy for family income? *British Educational Research Journal*, [e-journal] 36(4), pp.673 – 690. 10.1080/01411920903083111

Hong, Y., Chiu, C., Dweck, C., Lin, D. and Wan, W., 1999. Implicit theories, attributions, and coping: a meaning system approach. *Journal of Personality and Social Psychology*, [e-journal] 77(30) pp.588 – 599. [pdf] Available at: <implicit theories, attributions, and coping: a meaning system approach> [Accessed 30 April 2021].

Hoover-Dempsey, K. and Sandler, H., 1995. Parental involvement in children's education: Why does it make a difference? *Teachers College Record,* 97(2), pp.310 – 331. [pdf] Available at: <parental involvement in children's education: Why does it make a difference?> [Accessed 30 April 2021].

Hoover-Dempsey, K. and Sandler, H., 1997. Why do parents become involved in their children's education? *Review of Educational Research*, [e-journal] 67(1), pp.3-42 10.3102/00346543067001003

Hoover-Dempsey, K., Walker, J., Sandler, H., Whetsel, D., Green, C., Wilkins, A. and Closson, K., 2005. Why do parents become involved? research findings and implications. *The Elementary School Journal*, [e-journal] 106(2), pp.105 – 130. 10.1086/499194

Holford, A. and Rabe, B., 2020. *Impact of the universal infant free school meal policy.* [pdf] Available at: <impact of the universal infant free school meal policy> [Accessed 11 April 2021].

Houtenville, A. J. and Conway, K., 2008. Parental effort, school resources, and student achievement. *The Journal of Human Resources,* [e-journal] 43(2), pp.437 – 453. [pdf] Available at: <parental effort, school resources, and student achievement> [Accessed 30 April 2021].

Hutchinson, J., Reade, M. and Akhal, A., 2020. *Education in England: annual report.* [pdf] Available at: <education in England: annual report> [Accessed 10 March 2021].

Hutchinson, J., Robinson, D., Carr, D., Hunt, E., Crenna-Jennings, W. and Akhal, A., 2018. *Education in England: annual report 2018.* [pdf] Available at: <education in England: annual report 2018> [Accessed 10 March 2021].

Ilie, S., Sutherland, A. and Vignoles, A., 2017. Revisiting free school meal eligibility as a proxy for pupils socio-economic deprivation. *British Educational Research Journal*, [e-journal] 43(2), pp.253 –274. <British_Educational_Research_Journal-AM.pdf> [Accessed 25 February 2021].

Kettlewell, K., Sharp, C. and Eivers, E., 2020a. *International early learning and child well-being study (IELS): National summary report for England*. [pdf] Available at: <international early learning and child well-being study (IELS): national summary report for England> [Accessed 3 March 2021].

Kettlewell, K., Sharp, C., Lucas, M., Gambhir, G., Classick, R., Hope, C., Kollias, C. and Rutt, S., 2020b. *International early learning and child well-being study (IELS): National report for England*. [pdf] Available at: <international early learning and child well-being study (IELS): national report for England> [Accessed 3 March].

Kettlewell, K., Sharp, C. and Eivers, E., 2021. *International Early Learning and Child Well-being Study (IELS). Summary for Schools and Parents: England.* [pdf] Available at: <iels_england_summary_for_schools_and_parents.pdf> [Accessed 23 April 2021].

Law, J., Charlton, J. and Asmussen, K., 2017*. Language as a child wellbeing indicator.* [pdf] Available at: <language as a child wellbeing indicator> [Accessed 10 March 2021].

Lawson, D. and Mace, R., 2009. Trade-offs in modern parenting: a longitudinal study of sibling competition for parental care. *Evolution and Human Behavior,* [e-journal] 30(3), pp.170-183. 10.1098/rstb.2010.0297

Leonard, J., Garcia, A. and Schulz, L., 2019. Practice what you preach: How adults' actions, outcomes, and testimony affect preschoolers' persistence. *Child Development,* [e-journal] 2019, pp.1–18. 10.1111/cdev.13305

Lorah, J., 2020. Estimating a multilevel model with complex survey data: demonstration using TIMSS. *Journal of Modern Applied Statistical Methods*, [e-journal] 18(2), pp.2–20. [pdf] Available at: estimating a multilevel model with complex survey data: demonstration using TIMSS. [Accessed 30 April 2021].

Lucas, M., Hope, C. and Sharp, C., 2021. [Forthcoming]. *IELS thematic report: young children's physical development in England*. London: DfE.

McClelland, M.M., Geldhof, G.J., Cameron, C.E. and Wanless, S.B., 2015. Development and self-regulation. In: Overton, W.F., Molenaar, P.C.M. and Lerner, R.M, ed. *Handbook of Child Psychology and Developmental Science Theoretical Models of Human Development.* Seventh edn. Hoboken, NJ: John Wiley & Sons.

Melhuish, E., 2010. *Impact of the home learning environment on child cognitive development: Secondary analysis of data from 'growing up in Scotland'.* [pdf] Available at: <impact of the home learning environment on child cognitive development: secondary analysis of data from 'growing up in Scotland'> [Accessed 10 March 2021].

Melhuish, E. and Gardiner, J., 2018. *Study of early education and development (SEED): Impact study on early education use and child outcomes up to age four years* [pdf] Available: <study of early education and development (SEED): impact study on early education use and child outcomes up to age four years > [Accessed 27 April 2021].

Melhuish, E. and Gardiner, J., 2020. *Study of early education and development (SEED): Impact study on early education use and child outcomes up to age five years.* [pdf. Available at: <study of early education and development (SEED): impact study on early education use and child outcomes up to age five years>[Accessed 10 March 2021].

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A. and Wager T. The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn Psychol*, [e-journal] 41(1), pp.49-100. 10.1006/cogp.1999.0734.

Mokrova, I., O'Brien, M., Calkins, S.D., Leerkes, E.M. and Marcovitch, S., 2013. The role of persistence at preschool age in academic skills at kindergarten. *European Journal of Psychology of Education*, 28(4), pp.1495-1503. [pdf] Available at: https://core.ac.uk/download/pdf/345079057.pdf [Accessed 30 April 2021].

Montacute, R., 2020. *Social mobility and Covid-19: implications of the Covid-19 crisis for educational inequality* [pdf] Available at: <social mobility and Covid-19: implications of the Covid-19 crisis for educational inequality> [Accessed 15 February 2021].

Mullis, I.V.S. and Martin, M.O. eds., 2015. PIRLS 2016 assessment framework, 2nd edition. [pdf] Available at: <PIRLS 2016 assessment framework, 2nd edition> [Accessed 10 March 2021].

Nezlek, J. B., 2008. An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, [e-journal] 2(2), pp.842–860. 10.1111/j.1751-9004.2007.00059.x

Organisation for Economic Co-operation and Development (2020a). Early learning and child well-being: A study of five-year-olds in England, Estonia, and the United States [online]. https://www.oecd-ilibrary.org/education/early-learning-and-child-well-being_3990407f-en

Organisation for Economic Co-operation and Development 2020b. *Early learning and child well-being in England*. 10.1787/c235abf9-en

Organisation for Economic Co-operation and Development (2020c). Early learning and child well-being in Estonia. [online] 10.1787/15009dbe-en

Organisation for Economic Co-operation and Development (2020d). Early learning and child well-being in the United States. [online] 10.1787/198d8c99-en

Organisation for Economic Cooperation and Development, 2021. *International early learning and child well-being study: Technical report.* [pdf] Available at:<international early learning and child well-being study> [Accessed 10 March 2021].

Pascal, C., Bertram, T., Cullinane, C. and Holt-White, E. 2020. *Covid-19 and Social Mobility Impact Brief #4: Early Years*. [pdf] Available at: <covid-19 and social mobility impact brief number 4> [Accessed 29 April 2021].

Piek, J., Pitcher, T. and Hay, D., 1991. Motor coordination and kinaesthesis in boys with attention deficit- hyperactivity disorder. *Developmental Medicine and Child Neurology,* [e-journal] 41(3), pp.159-165 10.1111/j.1469-8749.1999.tb00575.x

Ofsted, 2020. *Covid-19 series: briefing on early years, October 2020.* [pdf] Available at: <Covid-19 series: briefing on early years, October 2020> [Accessed 29 April 2021].

Roberts, F., Mathers, S., Joshi, H., Sylva, K. and Jones, E., 2010. *Millennium cohort study briefing 8: childcare in the pre-school years*. [pdf] Available at: <millennium cohort study briefing 8: childcare in the pre-school years> [Accessed 10 March 2021].

Rose, S., Twist, L., Lord, P., Rutt, S., Badr, K., Hope, C. and Styles, B., 2021. *Impact of school closures on KS1: potential implications for practice in year 2.* [pdf] Available at: <impact of school closures on KS1: potential implications for practice in year 2> [Accessed 10 March 2021].

Rubin, D. B., 1987. *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc.

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R. and McClelland, M. M., 2017. Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, [e-journal] 109(8), pp.1120–1140. 10.1037/edu0000193.

Serpel, Z.N. and Esposito A. G., 2016. Development of executive functions: implications for educational policy and practice. *Policy Insights from the Behavioral and Brain Sciences*, [e-journal] 3(2), pp.203-210. 10.1177/2372732216654718.

Sellen, P., Huda, N., Gibson, S. and Oliver L., 2018. *Evaluation of universal infant free school meals.* [pdf] Available at: <evaluation of universal infant free school meals> [Accessed 11 April 2021].

Shuey, E. A. and Kankaraš, M., 2018. *The power and promise of early learning*. OECD Education Working Paper No. 186. [pdf] Available at: <the power and promise of early learning> [Accessed 27 April 2021].

Sim, M., Bélanger, J., Hocking, L., Dimova, S., Lakovidou, E., Janta, B. and Teager, W., 2018. *Teaching, pedagogy and practice in early years childcare: an evidence review.* [pdf] Available at: <teaching, pedagogy and practice in early years childcare: an evidence review> [Accessed 10 March 2021].

Siraj-Blatchford, I., Mayo, A., Mulhuish, E., Taggart, B., Sammons, P. and Sylva, K., 2011. *Performing against the odds: developmental trajectories of children in the EPPSE 3-16 study*. [pdf] Available at: <performing against the odds: developmental trajectories of children in the EPPSE 3-16 study> [Accessed 10 March 2021].

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G. and Carpenter, J. R., 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal*, [e-journal] 338, pp.157–160. 10.1136/bmj.b2393

Strand, S., Malmber, L. and Hall, J., 2016. *English as an additional language (EAL) and educational achievement in England: an analysis of the national pupil database.* [pdf] Available at: < English as an additional language (eal) and educational achievement in England: an analysis of the national pupil database> [Accessed 15 February 2021].

Sutherland, A., Ilie, S. and Vignoles A., 2015. *Factors associated with achievement: key stage 4.* [pdf] Available at: <factors associated with achievement: key stage 4> [Accessed 11 April 2021].

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B., 2004. *Effective pre-school education: the final report* (Technical Paper 12). [pdf] Available at: <effective pre-school education: the final report> [Accessed 27 April 2021].

Sylva, K., Barreau, S., Melhuish, E., Sammons, P., Siraj Blatchford I. and Taggart, B., 2007. In: British Educational Research Association Annual Conference, The effects of the home learning environment on children's developmental outcomes at age 7. 5–8 September Institute of Education, London.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. and Taggart, B., 2008. *Final report from the primary phase: pre-school, school and family influences on children's development during key stage* 2 (Age 7-11). [pdf] Available at: <final report from the primary phase: pre-school, school and family influences on children's development during key stage 2 (Age 7-11)> [Accessed 10 March 2021].

Sylva, K., Sammons, P., Melhuish, E. C., Siraj, I. and Taggart, B., 2020. Developing 21st century skills in early childhood: the contribution of process quality to self-regulation and pro-social behaviour. Zeitschrift für Erziehungswissenschaft, 23(3), pp 465–48. 10.1007%2Fs11618-020-00945-x

Tabachnick, B. G. and Fidell, L. S., 2013. *Using multivariate statistics*. 6th ed. Boston, MA: Pearson.

Taylor, C., 2018. The reliability of free school meal eligibility as a measure of socio-economic disadvantage: Evidence from the millennium cohort study in Wales. British J*ournal of Educational Studies*, [e-journal] 66(1), pp.29 –51. 10.1080/00071005.2017.1330464

Van Buuren, S. and Groothuis-Oudshoorn, K., 2011. Mice: multivariate imputation by chained equations in R. *Journal of Statistical Softwar*e, [e-journal] 45(3), pp.1-67. 10.18637/jss.v045.i03

Vitiello, V., Greenfield, D., Munis, P. and George, J., 2011. Cognitive flexibility, approaches to learning and academic school readiness in head start preschool children. *Early Education & Development*, [e-journal] 22(3), pp.388 – 410. 10.1080/10409289.2011.538366:

Wheater, R., Durbin, B., McNamara, S. and Classick, R., 2016. *Is mathematics education in England working for everyone?*. [pdf] Available: <is mathematics education in England working for everyone?> [Accessed 10 March 2021].

Wheater, R., Galvis, M. and Buzinin, R., 2020. *Parents' engagement in their children's learning in Northern Ireland*. [pdf] Available at: < parents' engagement in their children's learning in Northern Ireland> [Accessed 10 March 2021].

Appendix A: Methodology

Introduction

The analysis sought to investigate associations between predictor variables and outcomes of interest in a multivariate manner, where confounding factors, correlated with both the outcome and predictor variables, were controlled for. Due to the clustered nature of the data, where children were nested within schools, multi-level modelling was used to conduct the analysis. The general approach was to build explanatory models in a quasi-forward selection manner where predictor variables were added one at a time in order of the strength of their univariate relationship with the outcome variable as long as they increased the predictive power of the multivariate model. Missing data was imputed before running the analysis. The rationale for the analysis approach and methodology used is explained in detail below.

Addressing missing data

Missing data was a challenge in the IELS analysis, especially for the parent questionnaire, although other sources of data had relatively high response rates. There was both unit missing data, where there were entire cases missing (that is, all the data from a survey was missing for one child), and values missing for specific variables in returned cases. Missing data could potentially introduce bias to the analysis, if, for example, there were more missing data from children with certain characteristics (for example, those from lower socio-economic status backgrounds). If only the complete cases were analysed, this would bias the analysis towards the group of respondents who answered. In addition, it is not valid to compare two models with differing amounts of missing data when using MLwiN (Version 3.04, Charlton and others, 2019) as this distorts the likelihood ratio statistic used to evaluate model significance multi-level modelling, as implemented in the software MLwiN. Multi-level modelling had to be used in this study because of the nature of the data where children were clustered within schools. This means that for correct inferences to be made about school and pupil level effects being analysed simultaneously a technique such as multi-level modelling has to be used that accounts for this clustering (Nezlek, 2008). Mean imputation is sometimes used as a remedy to the problem of missing data, but this is not the best solution. This method biases the analysis so that the more missing data a variable has, the less statistical power the study has to detect an effect that is there in reality in the population (Donders and others, 2006). This issue necessitated the use of multiple imputation, a technique that uses available data to estimate the likely values of missing data. It is the gold standard for dealing with missing data that provides the most unbiased estimate of its real values. Depending on the available data there will be varying degrees of uncertainty on the values of missing data using multiple imputation. If the present values

on a variable can be predicted well from the other variables then there will be less uncertainty around its missing values, however if the prediction is poor then there will be greater uncertainty. Each imputed value in multiple imputation is a different random draw from a distribution of likely values so the degree of divergence of the imputed values between values reflects the degree of certainty of the estimates. The procedure works in a similar way to multiple regression, whereby a series of predictor variable are used to predict the missing values of an outcome variable in the dataset.

The IELS study had unit non-response rates of 32.71% for the parent questionnaire, 8.50% for part A of the teacher questionnaire (which asked about the teacher) and 10.24% for part B of the teacher questionnaire (which asked about the children). There was no unit missing data for the child assessment data. All 2,577 cases in the IELS dataset had complete assessment data across the 4 learning outcomes. In terms of individual variables, the economic and social status variables from the parent questionnaire had the highest non-response rates, especially household income (HHIncome) 40.1% non-response rate and Parent 1 Socio-economic Index (44.2%) (the non-response rate for Parent 2 Socio-economic Index was 43.4%). Full details of missing values for variables used in this report are given in Table 8.

Table 8: Missingness rates for IELS survey variables

Variable	Variable label	Percentage missing
ELADQ0501	The following questions ask/Was the child easily distracted during the assessments	1.05
ELADQ0502	The following questions ask/Did the child stay on task during the assessments	1.16
INVOLVE	Parent involvement	22.27
ELPAQ090100	ISCED 01 attendance – Age 0	32.71
ELPAQ090101	ISCED 01 attendance – Age 1	32.71
ELPAQ090102	ISCED 01 attendance – Age 2	32.71
ELPAQ090200	ISCED 02 attendance – Age 0	32.71
ELPAQ090201	ISCED 02 attendance – Age 1	32.71
ELPAQ090202	ISCED 02 attendance – Age 2	32.71
ELPAQ090203	ISCED 02 attendance – Age 3	32.71
ELPAQ090204	ISCED 02 attendance – Age 4	32.71
ELPAQ090205	ISCED 02 attendance – Age 5	32.71
ELPAQ09SVC00	Supervision and care attendance – Age 0	32.71

ELPAQ09SVC01	Supervision and care attendance – Age 1	32.71
ELPAQ09SVC02	Supervision and care attendance – Age 2	32.71
ELPAQ09SVC03	Supervision and care attendance – Age 3	32.71
ELPAQ09SVC04	Supervision and care attendance – Age 4	32.71
ELPAQ09SVC05	Supervision and care attendance – Age 5	32.71
ELPAQ1101	In a typical week/Read to your child from a book or e-book	33.33
ELPAQ1201	About how many children's books	33.45
ELPAQ0401	How often does your child/use computer, tablet	33.49
ICTDEV	Exposure to ICT devices	33.49
ELPAQ1108	In a typical week/Do things outside together like walking, ball games, swimming or cycling	33.53
ELPAQ1901	How old are Parents/Guardians – Parent/Guardian 1	33.57
IMMIG	Immigration background (dichotomous)	33.57
ELPAQ1111	In a typical week/Do educational activities	33.60
ELPAQ1113	In a typical week/Take your child to a special or paid activity outside of the home (e.g. sports clubs, dance, swimming lessons, language lessons)	33.64
ELPAQ1103	In a typical week/Draw pictures	33.72
ELPAQ1109	In a typical week/Do activities – letters	34.19
Stud_Lang	Student language most often spoken at home	34.50
PAREDYRS	Highest number of years of formal education for either parent	34.96
ELPAQ2001	What is the highest level of formal education – Parent/Guardian 1	35.35
SES	SES index score	36.52
ELPAQ1902	How old are Parents/Guardians – Parent/Guardian 2	37.95
ELPAQ0601	Has your child ever experienced low birth weight?	38.40
ELPAQ2101	What is your annual household income	40.12
HHIncome	Household income (amount in national currency)	40.12
ELPAQ2002	What is the highest level of formal education – Parent/Guardian 2	42.03

P2SEI	Socio-economic index of occupation score for Parent 2	43.42
P1SEI	Socio-economic index of occupation score for Parent 1	44.16

In the current study, the MICE R statistical package (van Buuren and Groothuis-Oudshoorn, 2011, version 3.11.0) was used to impute missing data, in which Predictive Mean Matching was used for imputing missing numeric variable values and Multinomial Logistic Regression for missing categorical variable values. Predictive Mean Matching 'copies' actual values of a variable and replaces its missing values with them, see van Buuren and others (2011) for more details. This is done by identifying a set of transplant value candidates that occurred when the predictor variable values were similar to those where a missing value occurred, and randomly selecting one of these candidates for each imputation, and repeating the process for each missing value. Multinomial logistic regression is a regression technique that aims to predict unordered categorical outcome data based on predictor variables that are associated with each outcome. Each outcome is associated with a certain likelihood base on the predictor values, and the multiple imputation method selects between the possible outcomes randomly, based on the weightings of the likelihoods. Using these methods, 5 different imputed datasets were created to establish standard errors for coefficient values of multi-level models.

Addressing non-response bias

Non-response bias is an issue for multiple imputation as it is based on information provided by the available data. Therefore, for the process to function appropriately data should be missing at random, so that there is no bias in 'missingness' after accounting for the predictor variables. For example, there should not be more boys missing than girls. To prevent this from happening, using as many predictor variables as possible is recommended (Sterne and others, 2009). Consequently, in the IELS study the full range of available variables was used.

Multi-level modelling – general approach

The general approach taken with the multi-level modelling in this study was to build explanatory models of outcome variables using a quasi-forward variable selection technique. This involved a two-stage process for each model. First a series of single-variable models was created, one for each variable in a selection, which were either under investigation themselves or were to be controlled for before a variable or interaction term of interest was added to investigate whether it significantly improved the predictive power of the model. The variables that were identified at this stage as being significant predictors of the dependent variable (learning outcome) were then selected as

candidates to be added into one multivariate model. This was done by first arranging them in descending order of significance based on their *t* statistics[32] and then considering each variable for inclusion in the model one at a time in that order. To be included in the model, a variable had to increase the predictive ability of the model. First the model started with only the constant, the mean value of the outcome variable, and then the most significant predictor from the single-variable models was tested to see if it increased the predictive power of the model. If it did, then the predictor would be added to the model and the next most significant predictor from the single-variable models would be tested, and so on. Likelihood ratio tests were used as the criterion to determine if a term led to a significant improvement in a model's statistical power, a statistic that reflects the amount of variance in the outcome measure explained by the model. To calculate this statistic the '-2 log likelihood' value for the model with a term being tested was subtracted from that for to a model with all previously added terms (which can be called the 'base model'). The resulting value was looked up on a chi square distribution for significance with the degrees of freedom being the number of extra terms added in the "test model" in addition to those in the base model. This was typically just the number of levels for a categorical variable added to a model or just a value of one for a continuous variable added, but the degrees of freedom were greater for some models where a series of interaction terms were included. All multi-level models were run with survey weights applied. Balanced repeat replicate weighting was not applied due to multi-level modelling being regarded as sufficient to account for the probability proportional to size sampling design of the IELS survey (Lorah, 2020).

Multi-level Modelling – implementation

Deprivation

For the deprivation analysis, in order to investigate whether school-level IDACI and child-level IDACI significantly interacted with each other, a base model to control for confounding main effects was built first using the general approach identified above. The next step was to test whether a 'school-level IDACI by child-level IDACI' interaction term significantly increased the predictive ability of the models. A set of models was also created where school-level IDACI was added as an interaction term to each of the main effects identified in the base model, and another set where pupil IDACI was added to each of the main effects in the base model.

To determine if economically deprived children had different predictors of performance compared with children from affluent backgrounds, the General Approach outlined above

[32] A t statistic is a standardised measure of significance calculated by dividing a regression slope coefficient by its standard error.

was used on a set of data comprised only of children in the lowest IDACI quartile and again separately on a set of data comprised of children in the highest IDACI quartile.

Multi-level modelling standard error calculation

Measurement error was quantified for both dependent variable values and missing independent variable values. For the dependent variables this information is available as a series of plausible values drawn from a distribution of likely values, and for the missing independent variable values this information is available as five different imputations drawn from a distribution of the likely missing values. This is an established method of running analysis for International Large Scale Assessments (ILSAs) and used by OECD to correctly estimate standard errors for their surveys (OECD, 2021) for more information. Using Rubin's rule (Rubin, 1987) this measurement error information can be combined with sampling error to produce an overall standard error for coefficient estimates. With this procedure, variation in point estimates from inferential models (for example, variation in coefficient values from regression models) is used to estimate measurement variance. In the current study plausible values were created by the IEA and were part of the dataset given to NFER before the multiple imputation of missing independent variable values was done. Therefore 25 unique datasets were created to avoid any bias that might be introduced with applying Rubin's rule with any particular combination of dependent variable plausible values and independent variable imputation values. Each dataset was a unique combination of independent variable imputations (imp) and dependent variable plausible values (pv). The 25 unique datasets were divided into 5 unique 'draws'. For example, draw 1 consisted of the following five data sets: i) imputation set 1 (imp1) and plausible values set 1 (pv1), ii) imputation set 2 (imp2) and plausible value set 2 (pv2), iii) imputation set 3 (imp3) and plausible value set 3 (pv3), (iv) imputation set 4 (imp4) and plausible value set 4 (pv4), (v) imputation set 5 (imp5) and plausible value set 5(pv5). Table 9 illustrates the 5 draws consisting of 5 unique combinations of an imputation set and a plausible value set.

Table 9: Illustration of the method used to estimate standard errors

Draw 1	Draw 2	Draw 3	Draw 4	Draw 5
imp1 & pv1	imp1 & pv2	imp1 & pv3	imp1 & pv4	imp1 & pv5
imp2 & pv2	imp2 & pv3	imp2 & pv4	imp2 & pv5	imp2 & pv1
imp3 & pv3	imp3 & pv4	imp3 & pv5	imp3 & pv1	imp3 & pv2
imp4 & pv4	imp4 & pv5	imp4 & pv1	imp4 & pv2	imp4 & pv3
imp5 & pv5	imp5 & pv1	imp5 & pv2	imp5 & pv3	imp5 & pv4

Rubin's rule (formula 1) was applied to each of the 5 draws identified above in table 9, to derive 5 different measurement variance estimates, and the mean of these was as the final measurement variance estimate for each statistical test. This was then added to the sampling variance provided by MLwiN and the square root of this value taken as the final standard error. Final coefficient values were determined as the mean coefficient values across each of the 25 datasets.

Figure 9: Formula 1. Measurement variance calculation (Rubin, 1987).

$$\left[\left(1+\frac{1}{P}\right)*\frac{\sum_{p=1}^{P}\left(\varepsilon_{0,p}-\bar{\varepsilon}_{0,P}\right)^2}{P-1}\right]$$

P is the number of imputations $\varepsilon_{0,p}$ is the point estimate for a particular dataset $\bar{\varepsilon}_{0,P}$ the mean of the point estimates across datasets.

Significance testing

In order to determine whether individual coefficients within models were significantly different from zero, t values were calculated by dividing the final coefficient values described above by their final standard errors and the corresponding p value for 4 degrees of freedom calculated. Four degrees of freedom were used as 5 different datasets were used for calculating variance estimates.

The significance of models was performed in such a way that base models were compared to test models based on the same dataset (for example, base model on imp 1 & pv 1 was compared to test model on imp 1 & pv 1) yielding 25 base model – test model pairs. Each pair had to have a significant log likelihood ratio in order for the tested term or set of terms to be deemed to significantly increase the predictive performance of the model.

Calculating differences in months

Table 10 provides information to assess the confidence intervals for all measures included in this thematic report, based on the relationship between age in months and each measure (rather than the relationships between age and learning outcomes within the multi-level models, which include the influence of other variables on the learning outcomes). For example, the monthly difference mean for the emergent literacy outcome was 7.29 with a standard error of 0.57 and a confidence interval of ± 1.12 (1.96 X 0.57). If

you resampled an infinite number of times from the population, each time getting a slightly different estimate of the population mean, and drew the confidence interval above around it, 95% of samples would contain the actual population mean. The confidence interval therefore provides an indication of the degree of uncertainty around mean estimates. Thus, for the emergent literacy outcome, every difference of 7.29 in the measure is equivalent to 1 month. In other words, a mean difference of 16.29 is equivalent to 2.24 (16.29 / 7.29 = 2.24) months.

It should be noted that confidence intervals also exist for mean differences. For measures where the mean difference between the youngest group of children (4 years and 11 months) and the oldest group of children (6 years and 0 months) was not statistically significant, the monthly mean difference was not estimated or reported (-).

Table 10: Average monthly difference for each outcome

Outcome	Monthly difference mean	Monthly difference S.E.	Confidence Interval
Emergent literacy	7.29	0.57	± 1.12
Emergent numeracy	9.80	0.57	± 1.12
Mental flexibility	6.28	0.75	± 1.47
Emotion identification	6.55	0.64	± 1.25

Source: OECD IELS England database matched to NPD

Overview of assumption checking

Before starting on the analysis, a statistical analysis plan was drawn up by the NFER team for review and agreement with independent experts in early childhood education research and early childhood education policy at the DfE. This set out 5 assumption checks to be performed on the data, as set out below.

1. **Linearity**

   Checking that the outcome variables followed a linear relationship with continuous predictor variables.

2. **Normality of Residuals**

   Making sure that the residuals, which are the differences between model predictions and the actual data, are approximately normally distributed.

3. **Homogeneity of Variance**

   Checking that residuals are homogenously distributed across the whole range of continuous predictor variables so that they are not systematically larger or smaller at any particular point on the scale.

4. **Independence of observations**

   Our multi-level models controlled for school-level clustering. Observations were then assumed to be independent at the child level and not clustered together by postcode or some other factor. The plan was to test this assumption by analysing whether the residuals were grouped together by any geographical data that might become available at the analysis stage.

5. **Multicollinearity**

   Predictor variables should not be correlated with each other when running regression-type models as this can lead to an underestimation of regression coefficient size, leading to a type-1 errors where real effects are missed. Principle components analysis (PCA) and correlation analyses on the predictor variables were planned to determine if multicollinearity of predictor variables was a problem.

Implementation of assumption checks

Due to the nature of the data some of the assumption checks were not possible and others had to be adjusted. It was not possible to check the Independence of observations check as no geographical data was available during the analysis to perform it, however this is often the case with survey studies and is typically not seen as a problem. The multicollinearity check had to be adjusted because the predictor variables were mostly categorical in nature, making a PCA analysis infeasible. Also the large number of predictors would have made interpretation of the PCA difficult. The PCA was therefore replaced with an analysis of the Variance Inflation Function (VIF), which calculates the degree to which multicollinearity between predictors is an issue for interpretation of a regression-type analysis.

Results of assumption checks

1. **Linearity**

   Visual inspection of bivariate scatter plots between continuous predictors and outcome variables showed that all of the relationships were approximately linear.

2. **Normality of Residuals**

   QQ plots of level-1 (pupil level) and level-2 (school level) residuals showed that they were approximately normally distributed.

3. Homogeneity of Variance

Visual inspection of scatter plots between continuous predictors and the outcome variables shows that residuals were approximately normally distributed.

4. Multicollinearity and assumption checking – VIF

When running multivariate analysis looking at the associations between many independent variables and a dependent variable, the issue of multicollinearity is a possible problem for any analysis. Given the data we are using it is fair to accept that there will be some correlation between background characteristics, but it is important to ensure that the correlation is not too high. Whilst the correlation matrix has been investigated for the analyses, an alternative method is to use the Variance Information Function (VIF). Running regression analysis through SPSS allows the VIF to be identified and whilst there would appear to be no universal agreement on what values are acceptable, a value of above 10 would indicate undesirable levels of collinearity (Hair and others, 1995). OLS regression models[33] were run for each of the learning outcomes as dependent variables and all relevant variables identified through the model selection process entered as independent variables. The following tables identify the resulting outputs.

Table 11: Emergent literacy - VIF

Variable	Variable label	VIF
SCHFSM	Proportion of Children on Free School Meals (Quartile)	2.99
PUPILFSM	Pupil level FSM (binary)	1.25
IDACIScore_15_AUT19	Child level IDACI Score	2.00
IDACIQUART	School IDACI Quartile	2.79
eth_ASIA	Asian Ethnicity	5.58
eth_BLAC	Black Ethnicity	3.57
eth_MIXD	Mixed Ethnicity	4.48
eth_WHIT	White Ethnicity	10.66
ITSEX	Child Sex	1.04
ITAGE	Child Age	1.78
YRGROUP	Child Year Group	1.74
ELPAQ09SVC05	Supervision and care attendance – Age 5	2.61
ELPAQ09SVC03	Supervision and care attendance – Age 3	5.12

[33] Analysis was completed on one of the imputed datasets and repeated on a second dataset to ensure consistency of findings.

ELPAQ09SVC02	Supervision and care attendance – Age 2	6.09
ELPAQ09SVC01	Supervision and care attendance – Age 1	4.82
ELPAQ09SVC00	Supervision and care attendance – Age 0	2.68
ELPAQ090101	ISCED 01 attendance – Age 1	2.24
ELPAQ090100	ISCED 01 attendance – Age 0	2.18
ELPAQ090205	ISCED 02 attendance – Age 5	1.32
ELPAQ090204	ISCED 02 attendance – Age 4	1.21
PAREDYRS	Highest number of years of formal education for either parent	5.90
SES	SES index score	3.48
IMMIG	Immigration background (dichotomous)	1.97
INVOLVE	Parent involvement	1.12
ELADQ0501	The following questions ask/Was the child easily distracted during the assessments	2.03
ELADQ0502	The following questions ask/Did the child stay on task during the assessments	2.04
ELPAQ2001	What is the highest level of formal education – Parent/Guardian 1	3.25
ELPAQ2002	What is the highest level of formal education – Parent/Guardian 2	2.34
ELPAQ1901	How old are Parents/Guardians – Parent/Guardian 1	1.20
SEN2	NPD secondary SEN type	1.15
ELPAQ0401	How often does your child/use computer, tablet	1.04
ELPAQ1201	About how many children’s books	1.66
TEACHPERSIST	Teacher rating of Child Persistence	1.10
ELPAQ1601	Number of siblings the child has	1.14
EAL	English as a second language	2.16
Overall effectiveness	OFSTED rating of school	1.10
ELPAQ1113	In a typical week/Take your child to a special or paid activity outside of the home (e.g. sports clubs, dance, swimming lessons, language lessons)	1.24
IELSI014083	In a typical week how many times do you do activities with your child involving words or sentences?	1.32

ELPAQ1101	In a typical week/Read to your child from a book or ebook	1.45
ELPAQ1104	In a typical week how many times do you have a back and forth conversation with your child?	1.14

Table 12: Emergent numeracy - VIF

Variable	Variable label	VIF
SCHFSM	Proportion of Children on Free School Meals (Quartile)	2.96
PUPILFSM	Pupil level FSM (binary)	1.24
IDACIScore_15_AUT19	Child level IDACI Score	1.96
IDACIQUART	School IDACI Quartile	2.75
ITAGE	Child Age	1.76
YRGROUP	Child Year Group	1.74
ELPAQ090101	ISCED 01 attendance - Age 1	1.20
ELPAQ090205	ISCED 02 attendance – Age 5	1.21
ELPAQ090204	ISCED 02 attendance – Age 4	1.19
PAREDYRS	Highest number of years of formal education for either parent	5.87
SES	SES index score	3.42
INVOLVE	Parent involvement	1.09
ELADQ0501	The following questions ask/Was the child easily distracted during the assessments	2.02
ELADQ0502	The following questions ask/Did the child stay on task during the assessments	2.04
ELPAQ2001	What is the highest level of formal education – Parent/Guardian 1	3.22
ELPAQ2002	What is the highest level of formal education – Parent/Guardian 2	2.31
ELPAQ1901	How old are Parents/Guardians – Parent/Guardian 1	1.65
ELPAQ1902	How old are Parents/Guardians – Parent/Guardian 2	1.51
SEN2	Secondary SEN type	1.14
ELPAQ0401	How often does your child/use computer, tablet	1.04

ELPAQ1201	About how many children's books	1.55
TEACHPERSIST	Teacher rating of Child Persistence	1.09
ELPAQ1601	Number of siblings the child has	1.12
EAL	English as a Second Language	1.21
Overall effectiveness	OFSTED rating of school effectiveness	1.09
ELPAQ1113	In a typical week/Take your child to a special or paid activity outside of the home (e.g. sports clubs, dance, swimming lessons, language lessons)	1.23
IELSI014083	In a typical week how often do you do activities involving words and sentences with your child?	1.27
ELPAQ1101	In a typical week/Read to your child from a book or ebook	1.44

Table 13: Emotion identification - VIF

Variable	Variable label	VIF
SCHFSM	Proportion of Children on Free School Meals (Quartile)	2.95
PUPILFSM	Pupil level FSM (binary)	1.22
IDACIScore_15_AUT19	Child level IDACI Score	1.95
IDACIQUART	School IDACI Quartile	2.70
ITSEX	Child Sex	1.04
ITAGE	Child Age	1.72
YRGROUP	Child Year Group	1.74
ELPAQ09SVC03	Supervision and care attendance – Age 3	4.25
ELPAQ09SVC02	Supervision and care attendance – Age 2	4.21
ELPAQ090204	ISCED 02 attendance – Age 4	1.11
PAREDYRS	Highest number of years of formal education for either parent	5.84
SES	SES index score	3.40
INVOLVE	Parent involvement	1.10
ELADQ0501	The following questions ask/Was the child easily distracted during the assessments	2.02

ELADQ0502	The following questions ask/Did the child stay on task during the assessments	2.04
ELPAQ2001	What is the highest level of formal education – Parent/Guardian 1	3.18
ELPAQ2002	What is the highest level of formal education – Parent/Guardian 2	2.28
ELPAQ1901	How old are Parents/Guardians – Parent/Guardian 1	1.19
SEN2	Secondary SEN Type from NPD database	1.14
ELPAQ1201	About how many children's books	1.48
TEACHPERSIST	Teacher rating of child persistence	1.09
ELPAQ1601	How many siblings does the child have	1.12
ELPAQ1113	In a typical week/Take your child to a special or paid activity outside of the home (e.g. sports clubs, dance, swimming lessons, language lessons)	1.22
IELSI014083	In a typical week how often do you do tasks involving words or sentences with your child?	1.31
ELPAQ1101	In a typical week/Read to your child from a book or ebook	1.44
ELPAQ1104	In a typical week how many times do you have a back and forth conversation with your child?	1.13

Table 14: Mental flexibility - VIF

Variable	Variable label	VIF
PUPILFSM	Proportion of Children on Free School Meals (Quartile)	1.20
IDACIScore_15_AUT19	Child level IDACI Score	1.75
IDACIQUART	School IDACI Quartile	1.72
ITAGE	Child Age	1.74
YRGROUP	Child Year Group	1.72
ELPAQ090205	ISCED 02 attendance – Age 5	1.09
PAREDYRS	Highest number of years of formal education for either parent	5.77
SES	SES index score	3.27
INVOLVE	Parent involvement	1.08
ELADQ0501	The following questions ask/Was the child easily distracted during the assessments	2.01
ELADQ0502	The following questions ask/Did the child stay on task during the assessments	2.02
ELPAQ2001	What is the highest level of formal education – Parent/Guardian 1	3.13
ELPAQ2002	What is the highest level of formal education – Parent/Guardian 2	2.27
SEN2	NPD secondary SEN	1.13
ELPAQ1201	About how many children's books	1.43
TEACHPERSIST	Teacher rating of Child Persistence	1.09
ELPAQ1113	In a typical week/Take your child to a special or paid activity outside of the home (e.g. sports clubs, dance, swimming lessons, language lessons)	1.19
IELSI014083	In a typical week how many times do you do activities with your child involving words or sentences?	1.26
ELPAQ1101	In a typical week/Read to your child from a book or e-book	1.42

Correlation analysis showed that none of the predictor variables has a Pearson's correlation above 0.9 as recommended by Tabachnick and Fidell (2013). Analysis of the VIF values for the predictor variables entered into the multivariate models showed that

there was only one value that was greater than 10 and given that this is due to its dichotomous nature rather than indicating a problem with multicollinearity, it is not considered an issue for our analysis (Allison, 2012). Together these findings indicate that multicollinearity was not an issue in this study.

Appendix B: Bivariate analysis on deprivation measures

Tables 15 to 18 present the results from the bivariate analysis which investigated the relationship between 4 of the deprivation measures (FSM, child-level IDACI, school-level IDACI and pupil level FSM) and each of the 4 IELS learning outcomes in turn (but not in combination with each other).

Table 15: Children eligible for FSM had lower development than children not eligible for FSM

Measure	FSM mean	Non FSM mean	Difference	Standard Error of difference	Equivalence of difference in months
Emergent literacy	476.30	520.05	-43.75*	7.07	6.00
Emergent numeracy	486.01	534.54	-48.52*	7.05	4.95
Mental flexibility	491.28	515.63	-24.35*	7.88	3.88
Emotion identification	471.22	501.29	-30.07*	6.45	4.59

* This is statistically significant ($p < 0.05$)

Source: IELS assessment of 2,462 children matched to NPD data

Table 16: Children with the highest levels of deprivation showed lower development than the least deprived children

Measure	Least Deprived IDACI quartile	Most Deprived IDACI quartile	Difference	Standard Error of difference	Equivalence of difference in months
Emergent literacy	548.30	483.61	64.69*	7.58	8.88
Emergent numeracy	558.62	497.57	61.04*	7.44	6.23
Mental flexibility	529.01	498.40	30.61*	7.81	4.87
Emotion identification	518.19	482.81	35.37*	7.32	5.40

* This is statistically significant ($p < 0.05$).

Source: IELS assessment of 2,454 children matched to NPD data

Table 17: Children attending schools with the most deprived populations showed lower development in all 4 learning outcomes

Measure	Lowest FSM quartile mean	Highest FSM quartile Mean	Difference	Standard Error of difference	Equivalence of difference in months
Emergent literacy	544.66	480.75	63.91*	7.70	8.77
Emergent numeracy	552.49	496.59	55.90*	7.92	5.70
Mental flexibility	520.72	497.61	23.11*	8.38	3.68
Emotion identification	509.91	477.40	32.51*	8.38	4.96

* This is statistically significant ($p <0.05$)

Source: IELS assessment of 2,375 children matched to NPD data.

Table 18: Children attending schools in the most deprived areas showed lower development in all 4 learning outcomes

Measure	Least Deprived IDACI quartile	Most Deprived IDACI quartile	Difference	Standard Error of difference	Equivalence of difference in months
Emergent literacy	544.90	478.92	65.98*	7.93	9.05
Emergent numeracy	553.34	493.15	60.18*	8.23	6.14
Mental flexibility	521.77	494.47	27.30*	7.54	4.35
Emotion identification	508.87	477.97	30.90*	8.34	4.72

* This is statistically significant ($p < 0.05$)

Source: IELS assessment of 2,520 children matched to NPD data

Appendix C: Final regression and effect sizes

Models for whole sample

Tables 19 to 22 show all the variables that were included in the final models run for all 5-year-olds in our sample. The percentage of variance explained by each model is calculated as follows. The initial stage of analysis is to run a 'base case' model which only contains the intercept and no explanatory variables. This model gives the variance at 2 levels: for schools and for children. The final model also provides these variances. By looking at the change in variance between the base and final models it can be determined how much child-level variance has been accounted for by the variables in that model

The following tables also present the amount of child variation[34] explained by each of the final models used in the IELS analysis (that is, reduction in pupil variance). The percentage of variance explained by each model is calculated as follows. The initial stage of analysis is to run a 'base case' model which only contains the intercept and no explanatory variables. This model gives the variance at 2 levels: for schools and for children. The final model also provides these variances. By looking at the change in variance between the base and final models it is possible to understand how much child level variance has been accounted for by the variables in that model. There is a degree of consistency in that the models explain more variation for emerging literacy and emerging numeracy than they do for emotional identification and mental flexibility.

Table 19: Emergent literacy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	1588.96	244.10	*	911.25	2266.68	-	-
Pupil variance	7911.70	361.79	*	6907.21	8916.19	-	-
Final model							
School variance	585.69	141.94	*	191.61	979.78	-	-
Pupil variance	4657.08	201.99	*	4096.28	5217.87	-	-

[34] The amount of variance explained is calculated by finding the difference between the level 1 variance for the base case model and the level 1 variance in the final model.

Reduction in pupil variance	-	-	-	-	**41.48%**	-	-
Constant	68.63	39.34	-	-40.59	177.85	-	-
English is main language	26.69	5.09	*	12.55	40.83	0.27	0.10
Not at all distracted	53.23	8.31	*	30.15	76.30	0.55	0.27
Not really distracted	38.53	8.14	*	15.92	61.13	0.40	0.18
Distracted to some degree	26.69	7.57	*	5.68	47.71	0.27	0.11
Not at all on task	-71.71	14.89	*	-113.06	-30.37	-0.74	-0.09
Not really on task	-41.62	9.47	*	-67.91	-15.33	-0.43	-0.10
On task to some degree	-17.14	5.08	*	-31.23	-3.05	-0.18	-0.08
Uses digital device at least 1/month	31.04	9.51	*	4.63	57.45	0.32	0.09
Use of digital device at least once a week	11.68	7.19	-	-8.27	31.63	0.12	0.06
Use of digital device Everyday	14.32	7.43	-	-6.31	34.96	0.15	0.07
Attends a paid for activity 1-2 days/week	10.77	4.11	-	-0.65	22.19	0.11	0.06
Attends a paid for activity 3-4 days in a week	11.60	7.31	-	-8.69	31.90	0.12	0.04
Attends a paid for activity 5-7	3.06	10.34	-	-25.65	31.77	0.03	0.01

days in a week							
Up to 10 books	3.69	14.26	-	-35.91	43.28	0.04	0.01
11-25 books	15.29	15.99	-	-29.12	59.69	0.16	0.05
26-50 books	2.98	14.57	-	-37.46	43.42	0.03	0.01
51-100 books	25.65	14.98	-	-15.95	67.25	0.26	0.12
More than 100 books	32.26	13.44	-	-5.06	69.58	0.33	0.15
School-level IDACI: 2nd least deprived quartile	24.97	6.20	*	7.77	42.17	0.26	0.11
School-level IDACI: 2nd most deprived quartile	8.66	7.36	-	-11.78	29.11	0.09	0.04
School-level IDACI: least deprived quartile	34.57	6.71	*	15.93	53.21	0.35	0.16
Strong/moderate parental involvement	17.13	4.05	*	5.88	28.38	0.18	0.08
Age	44.20	6.67	*	25.67	62.72	0.19	0.13
No SEN	44.99	6.35	*	27.37	62.61	0.46	0.15
SES index score	9.54	2.24	*	3.31	15.76	0.13	0.09
Always persistent	52.30	13.02	*	16.16	88.45	0.54	0.13
Often persistent	31.78	11.93	-	-1.34	64.90	0.33	0.14
Rarely persistent	-16.97	12.88	-	-52.73	18.80	-0.17	-0.06
Sometimes persistent	8.18	11.71	-	-24.34	40.70	0.08	0.04

Year 1	35.62	5.78	*	19.57	51.67	0.37	0.13

Source: IELS assessment of 2,577 children, age 5

Table 20: Emergent numeracy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	1376.91	217.32	*	773.55	1980.27	-	-
Pupil variance	8614.48	302.10	*	7775.73	9453.24	-	-
Final model							
School variance	708.65	165.07	*	250.36	1166.95	-	-
Pupil variance	4896.84	116.07	*	4574.59	5219.09	-	-
Reduction in pupil variance	-	-	-	-	**44.79%**	-	-
Constant	14.99	40.36	-	-97.08	127.05	-	-
Not at all distracted	54.58	9.51	*	28.18	80.97	0.55	0.27
Not really distracted	39.41	9.26	*	13.69	65.13	0.39	0.18
Distracted to some degree	25.45	8.16	*	2.78	48.11	0.25	0.11
Not at all on task	-83.96	15.49	*	-126.97	-40.95	-0.84	-0.10
Not really on task	-45.32	8.37	*	-68.55	-22.09	-0.45	-0.11
On task to some degree	-20.11	4.51	*	-32.63	-7.59	-0.20	-0.09
Up to 10 books	-1.96	16.07	-	-46.58	42.67	-0.02	-0.01

11-25 books	10.53	17.87	-	-39.07	60.13	0.11	0.04
26-50 books	9.18	18.17	-	-41.26	59.62	0.09	0.04
51-100 books	22.90	17.56	-	-25.86	71.66	0.23	0.10
More than 100 books	27.88	17.58	-	-20.93	76.69	0.28	0.12
Four siblings	-37.34	10.27	*	-65.85	-8.84	-0.37	-0.07
More than four Siblings	-11.52	14.48	-	-51.72	28.67	-0.12	-0.02
One Sibling	-9.01	5.18	-	-23.38	5.36	-0.09	-0.05
Three Siblings	-5.71	7.81	-	-27.39	15.97	-0.06	-0.02
Two Siblings	-10.29	5.51	-	-25.59	5.02	-0.10	-0.04
School IDACI: 2nd least deprived quartile	28.58	6.90	*	9.43	47.74	0.29	0.12
School IDACI: 2nd most deprived quartile	15.18	7.20	-	-4.82	35.19	0.15	0.06
School-level IDACI: least deprived quartile	38.08	7.30	*	17.82	58.33	0.38	0.17
Strong/moderate parental involvement	18.12	4.19	*	6.49	29.76	0.18	0.08
Age	63.11	6.96	*	43.80	82.42	0.26	0.19
No SEN	41.09	6.33	*	23.50	58.68	0.41	0.13
SES index score	10.28	2.39	*	3.66	16.90	0.14	0.10
Always persistent	63.81	11.59	*	31.62	96.00	0.64	0.15
Often persistent	39.44	10.35	*	10.71	68.17	0.39	0.17

Rarely persistent	-12.21	11.03	-	-42.84	18.43	-0.12	-0.04
Sometimes persistent	16.24	9.53	-	-10.21	42.70	0.16	0.08
Year 1	45.11	5.26	*	30.50	59.72	0.45	0.17

Source: IELS assessment of 2,577 children, age 5

Table 21: Emotion identification

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	660.18	173.06	*	179.71	1140.66	-	-
Pupil variance	8616.40	330.67	*	7698.32	9534.48	-	-
Final model							
School variance	547.14	123.55	*	204.12	890.16	-	-
Pupil variance	6511.10	226.43	*	5882.45	7139.76	-	-
Reduction in pupil variance	-	-	-	-	**25.12%**	-	-
Constant	2.20	47.55	-	-129.81	134.20	-	-
Not at all distracted	51.47	11.80	*	18.70	84.24	0.53	0.26
Not really distracted	37.90	10.18	*	9.65	66.15	0.39	0.18
Distracted to some degree	27.69	9.76	*	0.59	54.80	0.29	0.12
Not at all on task	-57.50	19.02	*	-110.31	-4.68	-0.60	-0.07
Not really on task	-32.59	10.03	*	-60.44	-4.73	-0.34	-0.08

On task to some degree	-6.32	6.42	-	-24.15	11.51	-0.07	-0.03
Have a back-and-forth conversation 1-2 days in a week	39.01	12.35	*	4.74	73.29	0.40	0.14
Have a back-and-forth conversation 3-4 days in a week	30.24	10.31	*	1.60	58.88	0.31	0.14
Have a back-and-forth conversation 5-7 days in a week	36.91	10.52	*	7.71	66.10	0.38	0.19
Up to 10 books	11.44	20.63	-	-45.83	68.70	0.12	0.04
11-25 books	30.24	23.10	-	-33.90	94.37	0.31	0.11
26-50 books	9.46	21.49	-	-50.20	69.12	0.10	0.04
51-100 books	24.55	20.04	-	-31.10	80.20	0.25	0.11
More than 100 books	22.89	21.16	-	-35.86	81.65	0.24	0.11
Strong/moderate parental involvement	13.45	4.16	*	1.91	24.99	0.14	0.06
Age	66.25	6.88	*	47.15	85.35	0.29	0.20
Boy	-22.25	4.52	*	-34.81	-9.69	-0.23	-0.12
School FSM: most deprived quartile	-22.94	7.87	*	-44.80	-1.08	-0.24	-0.10
School FSM: 2nd most deprived quartile	-9.80	7.59	-	-30.87	11.28	-0.10	-0.04

School FSM: 2nd least deprived quartile	-10.87	7.03	-	-30.39	8.64	-0.11	-0.05
No SEN	38.67	6.37	*	21.00	56.35	0.40	0.13
Always persistent	49.49	14.12	*	10.30	88.69	0.51	0.12
Often persistent	30.65	13.33	-	-6.36	67.67	0.32	0.14
Rarely persistent	4.00	12.39	-	-30.40	38.41	0.04	0.01
Sometimes persistent	19.61	12.06	-	-13.86	53.09	0.20	0.10

Source: IELS assessment of 2,577 children, age 5

Table 22: Mental flexibility

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	640.67	167.73	*	174.98	1106.36	-	-
Pupil variance	10761.77	344.90	*	9804.20	11719.34	-	-
Final model							
School variance	696.80	169.69	*	225.68	1167.92	-	-
Pupil variance	8270.33	262.53	*	7541.45	8999.21	-	-
Reduction in pupil variance	-	-	-	-	**23.21%**	-	-
Constant	243.08	54.42	*	92.00	394.16	-	-
Not at all distracted	48.29	11.26	*	17.02	79.56	0.45	0.22

Not really distracted	23.53	10.68	-	-6.11	53.17	0.22	0.10
Distracted to some degree	16.96	10.96	-	-13.47	47.40	0.16	0.07
Not at all on task	-85.20	22.22	*	-146.90	-23.50	-0.80	-0.09
Not really on task	-69.34	10.34	*	-98.04	-40.65	-0.65	-0.16
On task to some degree	-25.31	6.77	*	-44.10	-6.53	-0.24	-0.11
Doing Words and Sentences Activities 1-2 days in a week	10.19	9.76	-	-16.91	37.30	0.10	0.04
Doing Words and Sentences Activities 3-4 days in a week	18.45	9.16	-	-6.99	43.88	0.17	0.08
Doing Words and Sentences Activities 5-7 days in a week	24.35	9.24	-	-1.32	50.02	0.23	0.11
Age	29.63	10.03	*	1.79	57.46	0.11	0.08
No SEN	32.93	7.45	*	12.24	53.63	0.31	0.10
SES index score	7.05	2.48	*	0.17	13.94	0.09	0.06
Always persistent	45.12	13.10	*	8.75	81.49	0.42	0.10
Often persistent	35.41	11.10	*	4.59	66.22	0.33	0.15

Rarely persistent	-5.71	11.51	-	-37.68	26.25	-0.05	-0.02
Sometimes persistent	18.06	10.06	-	-9.86	45.99	0.17	0.08
Year 1	28.03	7.40	*	7.49	48.56	0.26	0.10

Source: IELS assessment of 2,577 children, age 5

Models for children living in the most deprived areas

Tables 23 to 26 show all the variables that were included in the final models run for children living in the most deprived areas (most deprived child-IDACI quartile). Figures 10 to 13 show the effect sizes of the factors statistically significantly associated with development across the 4 learning outcomes for children living in the most deprived areas. Where multiple categories for the same variable emerged as significant, the category with the largest effect size is shown. The diamonds depict the mean effect size, while the whiskers above and below the mean show the 95% confidence intervals, taking account of the standard error, meaning that there is 95% confidence that the true value lies within this range.

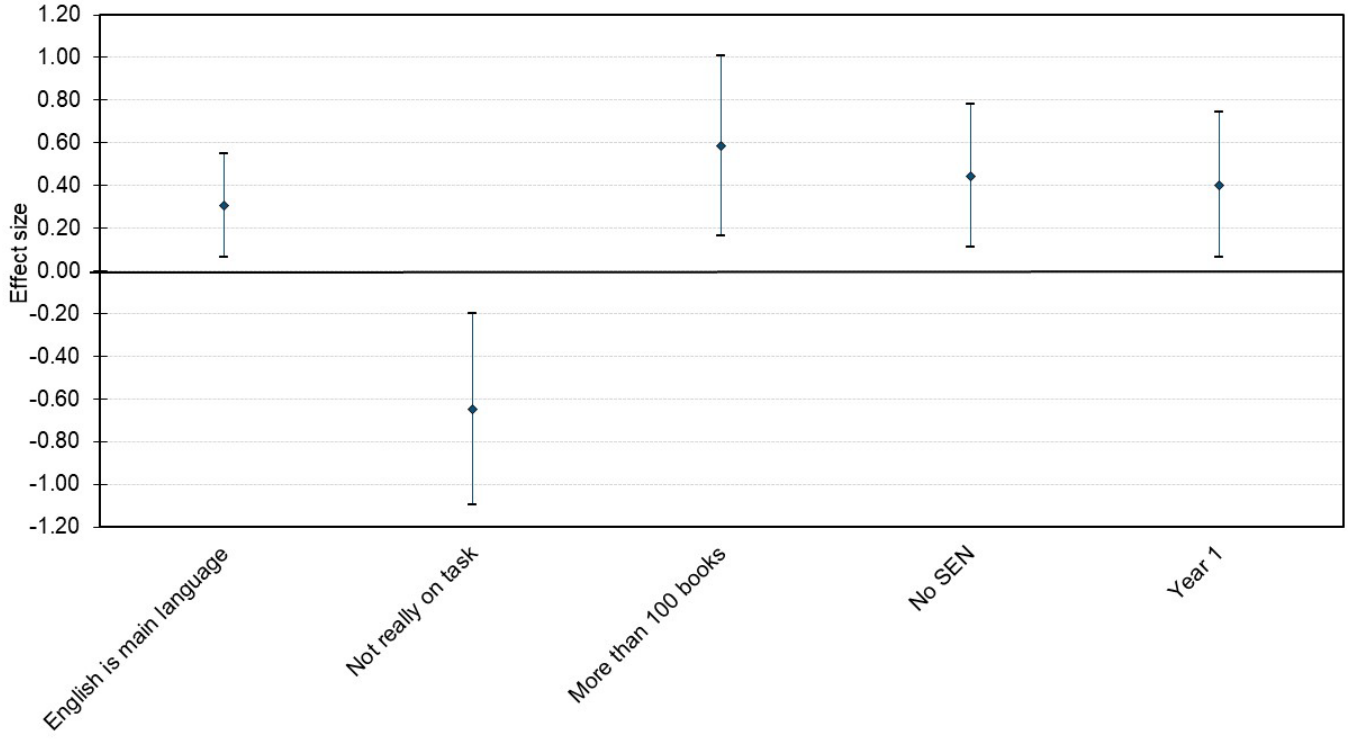
Table 23: High deprivation model - emergent literacy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	979.99	446.40	-	-259.40	2219.38	-	-
Pupil variance	8224.40	642.57	*	6440.35	10008.44	-	-
Final model							
School variance	582.92	359.10	-	-414.09	1579.93	-	-
Pupil variance	4959.05	495.68	*	3582.83	6335.27	-	-
Reduction in pupil variance	-	-	-	-	**36.70%**	-	-
Constant	145.31	73.19	-	-57.88	348.51	-	-

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
English is main language	29.37	8.37	*	6.12	52.61	0.31	0.13
Not at all distracted	42.35	18.45	-	-8.89	93.58	0.44	0.22
Not really distracted	24.14	17.37	-	-24.07	72.36	0.25	0.11
Distracted to some degree	21.61	16.65	-	-24.61	67.83	0.22	0.09
Not at all on task	-95.40	34.36	-	-190.80	0.01	-0.99	-0.14
Not really on task	-62.49	15.49	*	-105.50	-19.49	-0.65	-0.17
On task to some degree	-24.56	10.23	-	-52.95	3.83	-0.26	-0.12
11-25 books	17.41	12.29	-	-16.70	51.52	0.18	0.07
26-50 books	19.68	11.53	-	-12.32	51.68	0.20	0.09
51-100 books	35.36	12.43	*	0.86	69.87	0.37	0.15
More than 100 books	56.21	14.61	*	15.65	96.78	0.58	0.22
Age	33.85	14.34		-5.95	73.65	0.15	0.11
No SEN	42.81	11.61	*	10.58	75.05	0.45	0.15
Always persistent	71.43	26.50	-	-2.15	145.02	0.74	0.15
Often persistent	42.65	20.64	-	-14.64	99.95	0.44	0.19
Rarely persistent	-7.74	21.63	-	-67.79	52.31	-0.08	-0.03
Sometimes persistent	20.27	21.33	-	-38.96	79.50	0.21	0.11
Year 1	38.77	11.83	*	5.92	71.63	0.40	0.15

Source: IELS assessment of 644 children, age 5

Figure 10: Factors associated with emergent literacy for children living in the most deprived areas



Source: IELS assessment of 644 children, age 5

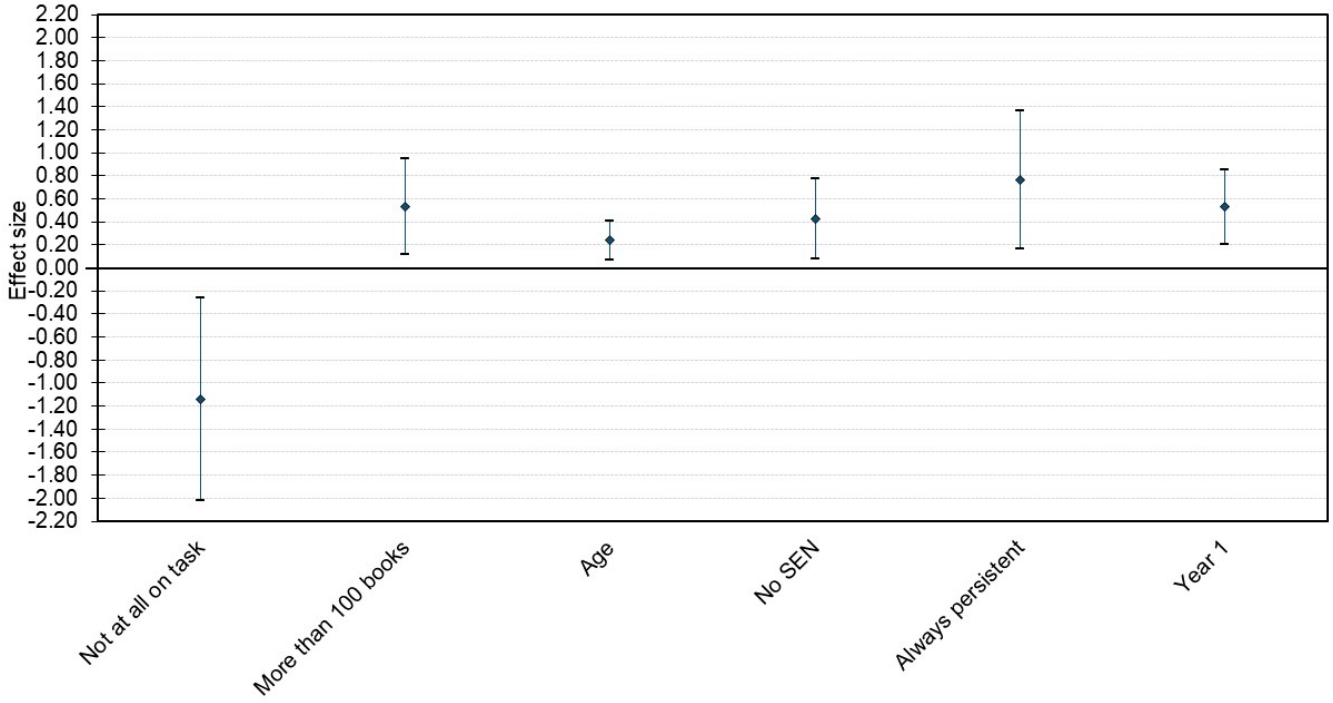
Table 24: High deprivation model - emergent numeracy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	679.03	371.72	-	-353.00	1711.07	-	-
Pupil variance	9012.29	638.29	*	7240.14	10784.44	-	-
Final model							
School variance	750.80	273.30	-	-7.99	1509.59	-	-
Pupil variance	4942.45	304.44	*	4097.21	5787.69	-	-

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Reduction in pupil variance	-	-	-	-	**46.33%**	-	-
Constant	46.43	72.67	-	-155.32	248.19	-	-
Not at all distracted	45.66	17.28	-	-2.31	93.63	0.46	0.23
Not really distracted	29.73	16.92	-	-17.24	76.69	0.30	0.14
Distracted to some degree	27.48	15.80	-	-16.37	71.33	0.28	0.12
Not at all on task	-112.33	31.28	*	-199.18	-25.49	-1.14	-0.16
Not really on task	-71.08	16.05	*	-115.64	-26.52	-0.72	-0.19
On task to some degree	-23.05	10.82	-	-53.08	6.98	-0.23	-0.11
11-25 books	16.38	11.60	-	-15.84	48.59	0.17	0.07
26-50 books	29.54	11.94	-	-3.62	62.70	0.30	0.13
51-100 books	34.38	11.88	*	1.41	67.36	0.35	0.14
More than 100 books	52.14	14.81	*	11.02	93.26	0.53	0.20
Age	55.22	14.21	*	15.76	94.68	0.24	0.17
No SEN	41.99	12.50	*	7.28	76.70	0.43	0.15
Always persistent	75.37	21.41	*	15.94	134.80	0.76	0.15
Often persistent	49.02	19.21	-	-4.31	102.35	0.50	0.21
Rarely persistent	1.36	18.36	-	-49.61	52.32	0.01	0.01
Sometimes persistent	19.36	16.22	-	-25.67	64.39	0.20	0.10
Year 1	52.21	11.53	*	20.20	84.22	0.53	0.20

Source: IELS assessment of 644 children, age 5

Figure 11: Factors associated with emergent numeracy for children living in the most deprived areas



Source: IELS assessment of 644 children, age 5

.

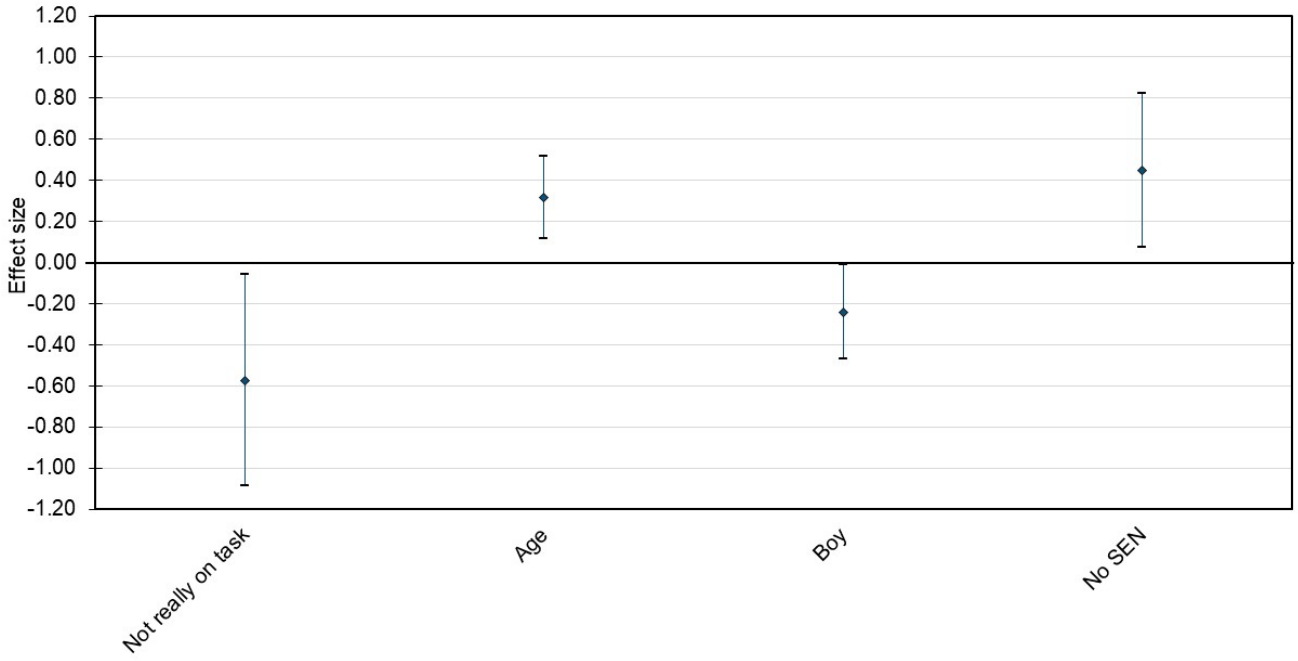
Table 25: High deprivation model - emotion identification

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	664.00	376.84	-	-382.25	1710.25	-	-
Pupil variance	9631.75	705.20	*	7673.82	11589.67	-	-
Final model							
School variance	739.96	321.22	-	-151.88	1631.80	-	-
Pupil variance	7082.66	554.12	*	5544.22	8621.11	-	-

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Reduction in pupil variance	-	-	-	-	**25.61%**	-	-
Constant	-2.75	102.46	-	-287.22	281.72	-	-
Not at all distracted	41.13	21.37	-	-18.20	100.47	0.40	0.20
Not really distracted	22.50	17.36	-	-25.68	70.69	0.22	0.10
Distracted to some degree	21.05	18.34	-	-29.87	71.96	0.21	0.09
Not at all on task	-67.46	48.40	-	-201.84	66.93	-0.66	-0.09
Not really on task	-58.19	18.77	*	-110.30	-6.09	-0.57	-0.15
On task to some degree	-9.55	12.41	-	-44.01	24.91	-0.09	-0.04
Age	75.94	17.26	*	28.03	123.86	0.32	0.22
Boy	-24.39	8.45	*	-47.84	-0.94	-0.24	-0.12
No SEN	45.51	13.75	*	7.34	83.67	0.45	0.16
Always persistent	58.49	25.47	-	-12.24	129.22	0.58	0.12
Often persistent	32.89	20.48	-	-23.96	89.74	0.32	0.14
Rarely persistent	1.61	21.51	-	-58.12	61.34	0.02	0.01
Sometimes persistent	22.15	19.50	-	-31.99	76.30	0.22	0.11

Source: IELS assessment of 644 children, age 5

Figure 12: Factors associated with emotion identification for children living in the most deprived areas



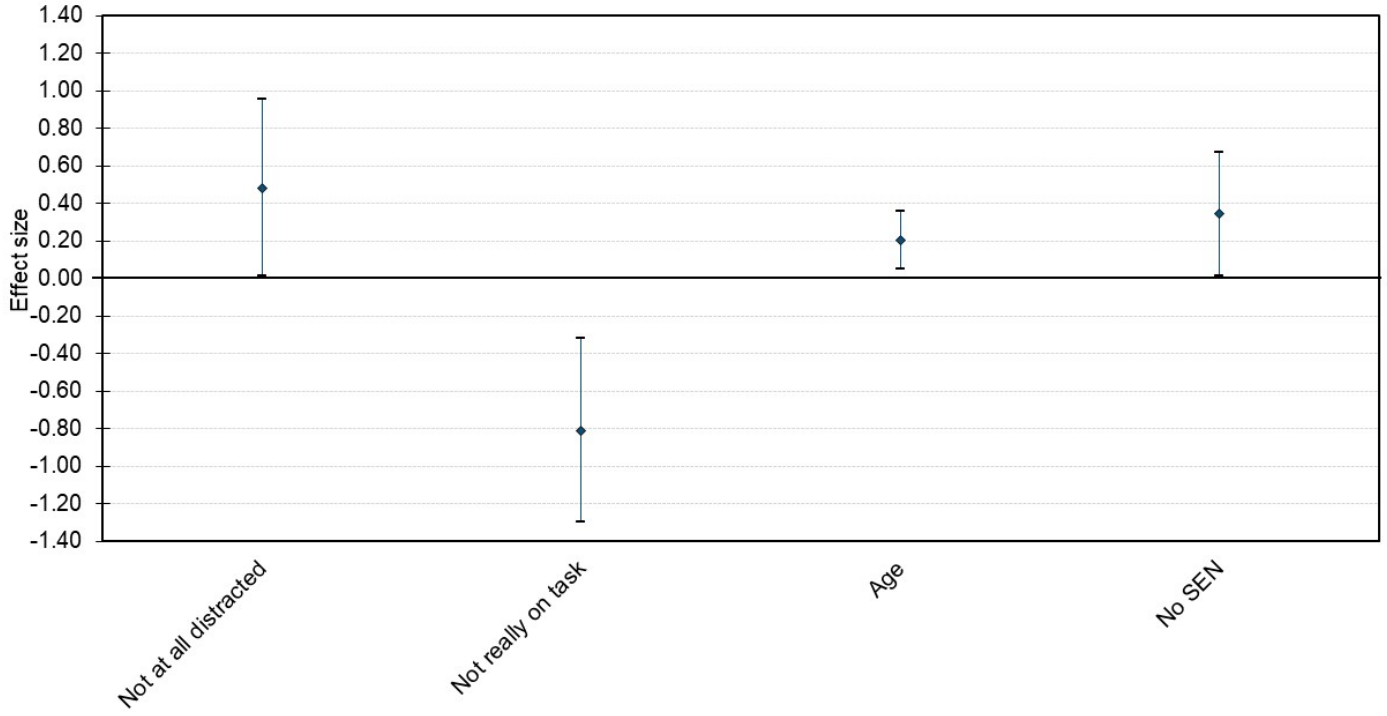
Source: IELS assessment of 644 children, age 5

Table 26: High deprivation model - mental flexibility

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	437.41	378.45	-	-613.30	1488.13	-	-
Pupil variance	11088.46	760.49	*	8977.03	13199.90	-	-
Final model							
School variance	701.67	397.38	-	-401.62	1804.96	-	-
Pupil variance	8517.00	563.93	*	6951.30	10082.69	-	-
Reduction in pupil variance	-	-	-	-	**23.62%**	-	-
Constant	166.22	81.27	-	-59.42	391.87	-	-
Not at all distracted	51.90	18.17	*	1.46	102.34	0.48	0.24
Not really distracted	30.80	17.80	-	-18.63	80.23	0.29	0.13
Distracted to some degree	17.78	18.07	-	-32.39	67.94	0.17	0.07
Not at all on task	-88.55	45.16	-	-213.92	36.83	-0.82	-0.12
Not really on task	-87.06	19.00	*	-139.81	-34.31	-0.81	-0.21
On task to some degree	-27.78	13.70	-	-65.81	10.25	-0.26	-0.12
Age	51.67	14.12	*	12.46	90.88	0.20	0.14
No SEN	36.78	12.65	*	1.66	71.89	0.34	0.12

Source: IELS assessment of 644 children, age 5

Figure 13: Factors associated with mental flexibility for children living in the most deprived areas



Source: IELS assessment of 644 children, age 5

Models for children living in the least deprived areas

Tables 27 to 30 show all the variables that were included in the final models run for children living in the least deprived areas (the least deprived IDACI quartile). Figures 14 to 17 show the effect sizes of the factors statistically significantly associated with development across the 4 learning outcomes for children living in the least deprived areas. Where multiple categories for the same variable emerged as significant, the category with the largest effect size is shown. The diamonds depict the mean effect size, while the whiskers above and below the mean show the 95% confidence intervals, taking account of the standard error, meaning that there is 95% confidence that the true value lies within this range.

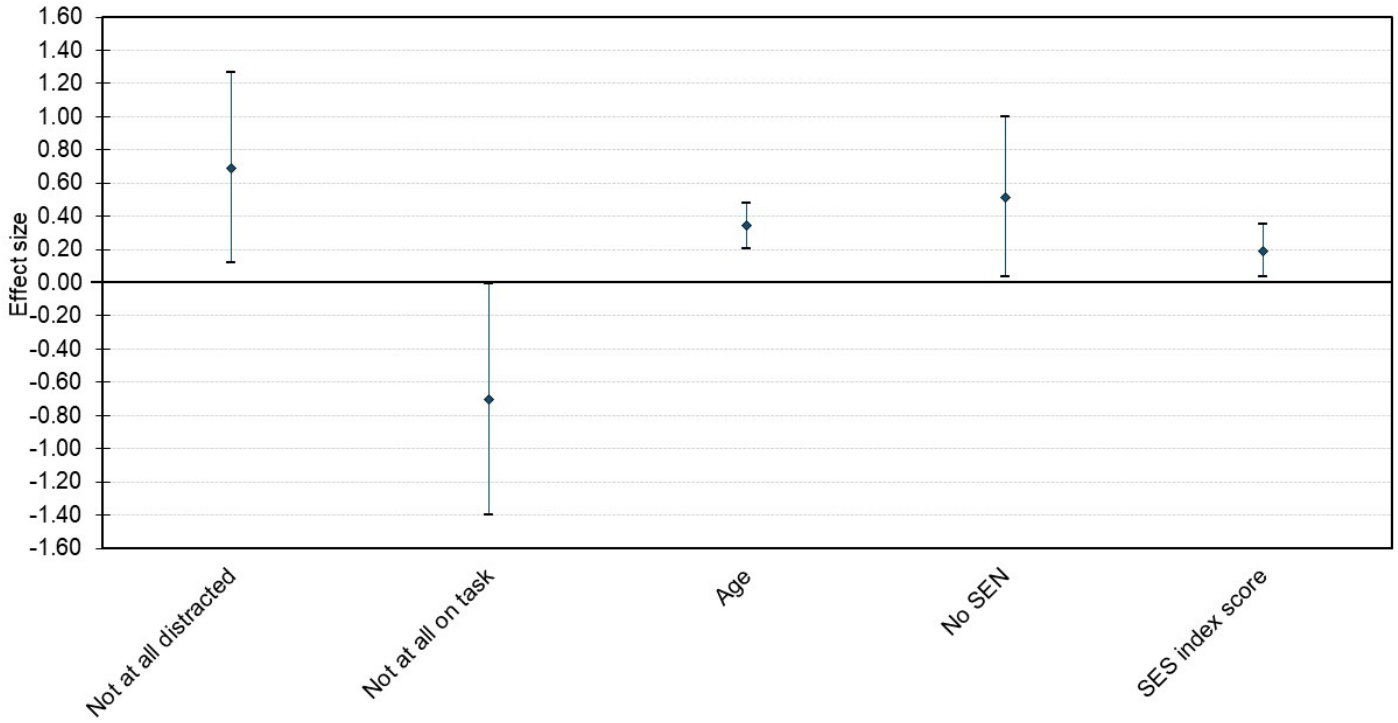
Table 27: Low deprivation model - emergent literacy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	881.42	310.07	*	20.54	1742.30	-	-
Pupil variance	7158.40	735.72	*	5115.74	9201.06	-	-
Final model							
School variance	732.99	240.85	*	64.30	1401.68	-	-
Pupil variance	4296.97	388.98	*	3216.99	5376.94	-	-
Reduction in pupil variance	-	-	-	-	**41.56%**	-	-
Constant	33.57	65.70	-	-148.85	215.98	-	-
Not at all distracted	62.11	18.50	*	10.74	113.48	0.69	0.34
Not really distracted	48.47	18.14	-	-1.89	98.83	0.54	0.24
Distracted to some degree	36.02	17.53	-	-12.65	84.69	0.40	0.16
Not at all on task	-62.97	22.56	*	-125.61	-0.34	-0.70	-0.09
Not really on task	-49.67	18.21	-	-100.23	0.89	-0.55	-0.12
On task to some degree	-21.55	8.27	-	-44.52	1.42	-0.24	-0.11
11-25 books	11.94	24.28	-	-55.48	79.35	0.13	0.03
26-50 books	3.43	22.69	-	-59.57	66.43	0.04	0.01
51-100 books	24.62	22.44	-	-37.69	86.92	0.27	0.13
More than 100 books	36.31	22.31	-	-25.63	98.25	0.40	0.20
Age	73.38	10.56	*	44.06	102.69	0.34	0.24

No SEN	46.28	15.62	*	2.92	89.64	0.52	0.14
SES index score	15.20	4.54	*	2.61	27.79	0.19	0.14
Always persistent	21.37	21.67	-	-38.79	81.54	0.24	0.05
Often persistent	18.86	19.80	-	-36.10	73.82	0.21	0.10
Rarely persistent	-31.20	20.51	-	-88.15	25.75	-0.35	-0.11
Sometimes persistent	-2.85	18.04	-	-52.93	47.24	-0.03	-0.02

Source: IELS assessment of 644 children, age 5

Figure 14: Factors associated with emergent literacy for children living in the least deprived areas



Source: IELS assessment of 644 children, age 5

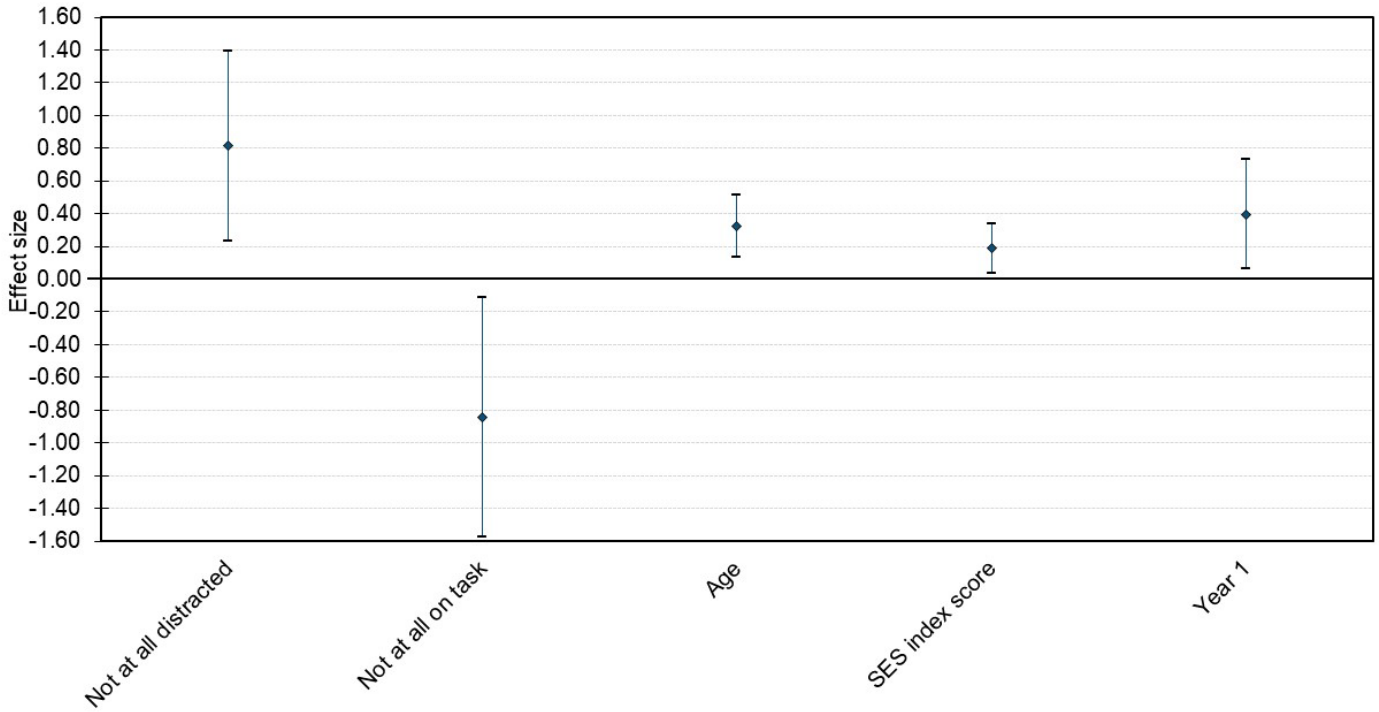
Table 28: Low deprivation model - emergent numeracy

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	703.50	313.79	-	-167.71	1574.70	-	-
Pupil variance	7745.40	698.05	*	5807.32	9683.47	-	-
Final model							
School variance	625.26	214.83	*	28.79	1221.72	-	-
Pupil variance	4490.27	436.03	*	3279.68	5700.86	-	-
Reduction in pupil variance	-	-	-	-	**41.13%**	-	-
Constant	54.02	84.18	-	-179.68	287.73	-	-
Not at all distracted	74.79	19.23	*	21.40	128.18	0.81	0.41
Not really distracted	58.43	17.81	*	8.97	107.89	0.64	0.29
Distracted to some degree	49.44	17.56	*	0.70	98.19	0.54	0.22
Not at all on task	-77.50	24.26	*	-144.84	-10.15	-0.84	-0.11
Not really on task	-43.80	18.75	-	-95.85	8.25	-0.48	-0.10
On task to some degree	-20.73	9.58	-	-47.34	5.88	-0.23	-0.10
ISCED 02 attendance - Age 4 <20 hrs	-7.20	17.25	-	-55.10	40.70	-0.08	-0.03
ISCED 02 attendance - Age 4 >20 hrs	15.57	10.42	-	-13.37	44.51	0.17	0.07

Strong/moderate parental involvement	18.81	7.60	-	-2.29	39.91	0.20	0.09
Age	70.78	14.94	*	29.31	112.25	0.32	0.23
SES index score	15.11	4.44	*	2.79	27.44	0.19	0.13
Always persistent	27.85	22.26	-	-33.95	89.64	0.30	0.07
Often persistent	22.33	16.56	-	-23.65	68.31	0.24	0.11
Rarely persistent	-36.64	19.30	-	-90.23	16.95	-0.40	-0.13
Sometimes persistent	3.17	14.92	-	-38.25	44.58	0.03	0.02
Year 1	36.54	11.10	*	5.72	67.36	0.40	0.15

Source: IELS assessment of 644 children, age 5

Figure 15: Factors associated with emergent numeracy for children living in the least deprived areas



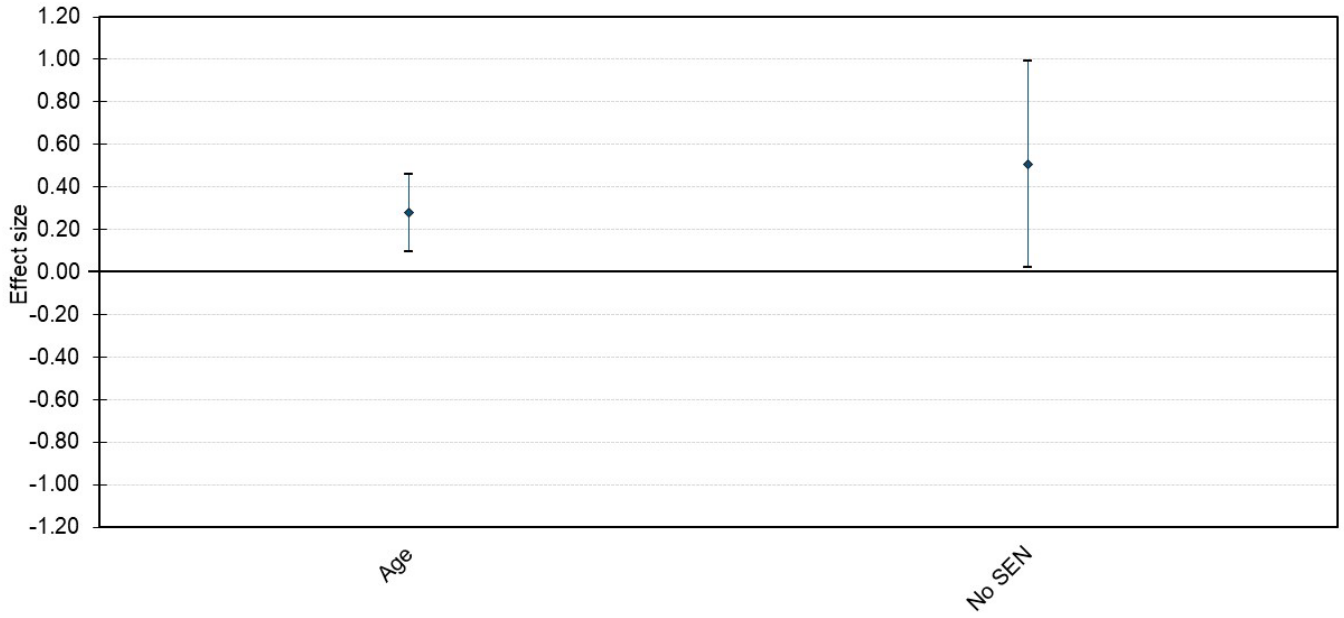
Source: IELS assessment of 644 children, age 5

Table 29: Low deprivation model - emotion identification

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base case							
School variance	597.91	345.07	-	-360.14	1555.97	-	-
Pupil variance	7717.75	745.32	*	5648.45	9787.06	-	-
Final model							
School variance	475.08	260.50	-	-248.18	1198.33	-	-
Pupil variance	6227.51	546.63	*	4709.84	7745.18	-	-
Reduction in pupil variance	-	-	-	-	**20.86%**	-	-
Constant	113.00	84.91	-	-122.74	348.75	-	-
Not at all distracted	48.33	22.12	-	-13.09	109.76	0.53	0.26
Not really distracted	35.78	20.61	-	-21.43	93.00	0.39	0.18
Distracted to some degree	17.34	21.16	-	-41.42	76.09	0.19	0.08
Not at all on task	-69.79	36.85	-	-172.11	32.54	-0.77	-0.10
Not really on task	-59.78	23.76	-	-125.74	6.17	-0.66	-0.14
On task to some degree	-12.55	12.69	-	-47.79	22.70	-0.14	-0.06
Age	60.57	14.34	*	20.75	100.40	0.28	0.20
No SEN	46.19	15.96	*	1.88	90.51	0.51	0.14

Source: IELS assessment of 644 children, age 5

Figure 16: Factors associated with emotion identification for children living in the least deprived areas



Source: IELS assessment of 644 children, age 5

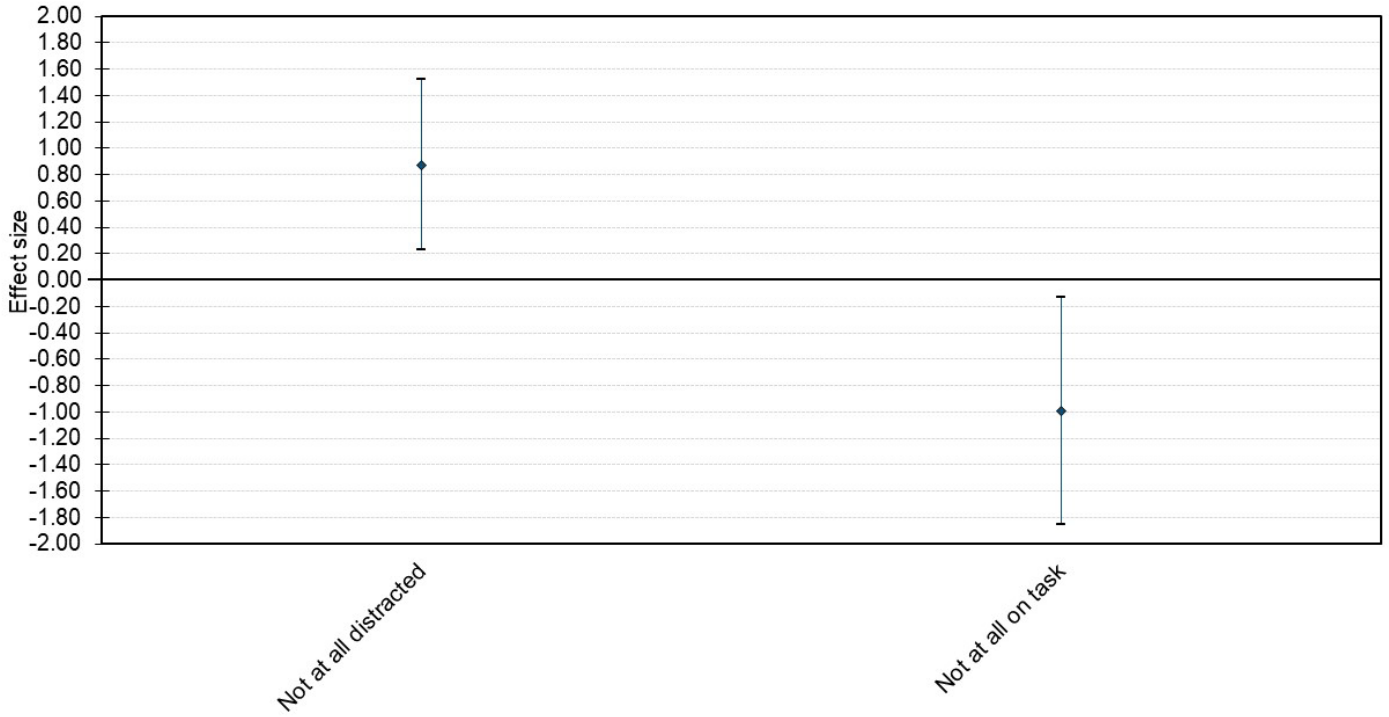
Table 30: Low deprivation model - mental flexibility

Parameter	Estimate	Standard error	Sig.	Min.	Max.	Effect Size Mean	Standardised Coefficient
Base Case							
School variance	819.86	362.04	-	-185.30	1825.02	-	-
Pupil variance	10472.07	752.31	*	8383.34	12560.79	-	-
Final model							
School variance	626.14	340.23	-	-318.47	1570.75	-	-
Pupil variance	9023.60	647.40	*	7226.16	10821.05	-	-
Reduction in pupil variance	-	-	-	-	**13.85%**	-	-

Constant	473.33	23.74	*	407.43	539.23	-	-
Not at all distracted	93.03	24.68	*	24.49	161.56	0.87	0.44
Not really distracted	66.71	22.60	*	3.97	129.45	0.63	0.28
Distracted to some degree	55.96	24.31	-	-11.54	123.45	0.53	0.22
Not at all on task	-105.98	32.97	*	-197.51	-14.44	-1.00	-0.13
Not really on task	-52.88	26.05	-	-125.20	19.44	-0.50	-0.11
On task to some degree	-34.69	13.48	-	-72.12	2.74	-0.33	-0.15

Source: IELS assessment of 644 children, age 5

Figure 17: Factors associated with mental flexibility for children living in the least deprived areas



Source: IELS assessment of 644 children, age 5

NFER was contracted to carry out IELS in England on behalf of the Department for Education (DfE) and this report includes analysis of pupil administrative data from the DfE's national pupil database (NPD). However the views expressed in this report are the authors' and do not necessarily reflect those of the DfE.

Please note that this work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Department for Education

ISBN: 978-1-83870-274-8

For any enquiries regarding this publication, contact us at:
EY.AnalysisAndResearch@education.gov.uk or www.education.gov.uk/contactus

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