Article:
Atkinson, AL, Hill, LJB, Pettinger, KJ et al. (2022) Can holistic school readiness evaluations predict academic achievement and special educational needs status? Evidence from the Early Years Foundation Stage Profile. Learning and Instruction, 77. 101537. ISSN: 0959-4752

https://doi.org/10.1016/j.learninstruc.2021.101537

Can holistic school readiness evaluations predict academic achievement and special educational needs status? Evidence from the Early Years Foundation Stage Profile

Abstract

This study examined the predictive validity of holistic school readiness evaluations using the 'good level of development' outcome from the Early Years Foundation Stage Profile (EYFSP). The EYFSP assesses a range of abilities at school entry including academic, language, socio-emotional, and motor skills. In particular, we examined whether the assessment predicted reading, writing, maths, and science ability two years later and future special educational needs (SEN) status (*N*=3,739-5,768). Children who reached a good level of development had higher odds of performing at expected (vs. below expected) levels on later academic assessments. This was particularly true for children with SEN. Reaching a good level of development also increased the odds of performing at above expected (vs. expected) levels on the academic assessments and lowered the odds of requiring SEN support. This demonstrates that holistic school readiness evaluations are powerful tools that can identify 'at risk' children.

Keywords: educational attainment; special educational needs; school readiness; primary education

1. Introduction

Not all children enter school exhibiting the same skills or readiness to learn (Claessens & Engel, 2013), with a myriad of contributing factors (e.g. Duncan & Magnuson, 2011; Isaacs & Magnuson, 2011; Pettinger et al., 2019; Ziol-Guest & McKenna, 2013). School readiness is therefore a topic of great interest for both researchers and policymakers (Davies et al., 2016). Schol readiness is broadly defined with reference to skills possessed by a child at the start of formal education that are critical for later academic success (Aiona, 2005; Snow, 2006), There is, however, disagreement regarding the precise definition, and the abilities that should be assessed as part of school readiness evaluations (Ackerman & Barnett, 2005; Dockett & Perry, 2002). Early academic skills are important predictors of later educational performance (Claessens et al., 2009; Claessens & Engel, 2013; Duncan et al., 2007; Stormont et al., 2015), leading some commentators to suggest that early years educators should focus primarily on ensuring children have strong academic abilities at school entry (e.g. Claessens et al., 2009). This is also the predominant position of parents, who generally view school readiness as contingent on academic skills (Barbarin et al., 2008). However, in direct contrast to this, teachers have suggested that non-academic abilities (e.g. the capacity to communicate well) are particularly important at school entry (Heaviside & Farris, 1993; Lin et al., 2003). Supporting this view, there is clear evidence that more general abilities, such as language (e.g. Duncan et al., 2007; Pagani et al., 2010) and motor skills (e.g. Grissmer et al., 2010; Pagani et al., 2010) are predictive of later educational performance.

A more moderate view of these extremes is that educators should take a holistic approach, focusing on developing both academic and broader developmental skills (Darling-Hammond & Cook-Harvey, 2018; Diamond, 2010). Proponents of these views suggest that

SEN = special educational needs; GCSEs = General Certificate of Secondary Education; OR = Odds Ratio; CI = Confidence Interval

this approach is likely to enhance children's physical health, emotional wellbeing, and overall development (Darling-Hammond & Cook-Harvey, 2018; Diamond, 2010). Diamond (2010) also argues that this approach is the optimal way of enhancing children's academic achievement, stating that *"If we want the best academic outcomes, the most efficient and cost-effective route to achieve that is, counterintuitively, not to narrowly focus on academics, but to also address children's social, emotional, and physical development."* (Diamond, 2010, p. 780). Holistic perspectives have focused on education more broadly but these ideas could be applied to school readiness: to ensure children are ready for the demands of classroom learning, they should have basic academic knowledge, good language skills, and be well-developed socially, emotionally and physically (Davies et al., 2016). Children who arrive at school with these abilities are more likely to exhibit a range of foundational skills that are important for classroom learning, such as the ability to follow lessons, build positive relationships with teachers, co-ordinate their movements well, and follow instructions (e.g. Lopes et al., 2013; Davies et al., 2016; Dunham et al., 2020; Jaroslawska et al., 2016; Parker et al., 2004; Ryan et al., 1994; Waterman et al., 2017). Such perspectives would therefore suggest that school readiness should not be viewed in terms of isolated skills, but should instead be considered as a multidimensional construct, comprising both academic and non-academic abilities (e.g. vocabulary, motor skills, socio-emotional skills). Indeed, this multidimensional approach has been adopted by some researchers, governments, and policy advisors across the world, who view school readiness as being contingent on both academic and non-academic abilities (e.g. Brinkman et al., 2013; Copple, 1997; Forget-Dubois et al., 2007; Standards and Testing Agency, 2017; Head Start, 2020).

But is performance on holistic school readiness evaluations, which assess both academic and non-academic skills, predictive of later educational performance? Relatively little research has investigated this to date. In some studies that have explored this, children

have been categorized based on their strengths and weaknesses across several domains (e.g. social/emotional, language, cognition, and health; Hair et al., 2006). In these studies, children showing a 'comprehensively positive profile', with above average performance in all domains, performed the best on later academic assessments (Hair et al., 2006; Pan et al., 2019). In contrast, those who performed below average across domains exhibited the poorest educational achievement (Hair et al., 2006). However, by assessing an individual relative to the mean, it makes this a comparative rather than standardized evaluation and thus impossible for all children to be characterized as showing a 'comprehensively positive profile' (and therefore being school ready across domains).

A more intuitive way to assess school readiness would be to generate a series of statements that describes domain-specific skills that children should have at school entry. Children could then be categorized as being school ready if they have acquired these skills. Indeed, this approach has been taken to track school readiness at a population-level in several countries (e.g. Canada, Australia). Although only a limited number of studies have investigated the predictive validity of these assessments, the findings thus far have been promising (Brinkman et al., 2013; Forget-Dubois et al., 2007). For instance, Brinkman et al. (2013) found that the Australian Early Development Index conducted at aged five predicted educational performance at eight, 10 and 12 years of age. However, as studies to date have used variants of the same assessment (the Early Development Index), further research is needed to investigate whether other holistic school readiness measures can predict later educational outcomes. This would allow previous work using the Early Development Index to be generalized, demonstrating that holistic school readiness evaluations are predictive of later academic achievement.

A statutory assessment that is broadly similar to the Early Development Index is conducted in England (the Early Years Foundation Stage Profile). This is a standardized

evaluation of children's abilities conducted in the year before the start of formal, classroom-based education (Reception; Standards and Testing Agency, 2017)[2]. The assessment is detailed and holistic, evaluating children against early learning goals that cover three prime areas of learning ('communication and language development', 'physical development' and 'personal, social and emotional development') and four specific areas of learning ('literacy', 'numeracy', 'understanding of the world' and 'expressive arts and design'). For each early learning goal, children are rated as having skills that are either 'emerging', 'expected' or 'exceeding'. Children are then given a total score. A school readiness measure (the 'good level of development' outcome; Office for Standards in Education, Children's Services and Skills, 2014; Pettinger et al., 2019; Public Health England, 2019) is also calculated. Children are rated as having reached a good level of development only if they are performing at 'expected' levels or 'exceeding' expected levels on all of the early learning goals that assess 'communication and language development', 'physical development', 'personal, social and emotional development', 'literacy', and 'numeracy'. As well as being used to track population-level school readiness, the assessment also provides Key Stage 1 teachers with pupil-level data, to enable them to plan lessons that meet the needs of all their pupils (Standards and Testing Agency, 2017).

Despite the Early Years Foundation Stage Profile being conducted on over 7.5 million children since its introduction in 2008 (Department for Education, 2012; 2019a), very little research has investigated the extent to which this assessment can predict later academic achievement. To the best of our knowledge, only one study has investigated this thus far

[2] In England, most children aged 4-5 years of age complete 'Reception'. As part of this academic year, children are exposed to key subject areas, such as reading and maths, although there is also a strong focus on play (Office for Standards in Education, Children's Services and Skills [Ofsted], 2017). While Reception is sometimes regarded as the first year of school (Ofsted, 2017), it is not compulsory for all children (depending on their birth month; UK Government, N.D.a) and fits within the early years foundation stage (UK Government, N.D.b). Following Reception, children in mainstream education then enter Key Stage 1, which is the first stage of compulsory classroom-based education.

(Treadaway, 2019). Treadaway (2019) reported a series of analyses which demonstrated that the total score on the Early Years Foundation Stage Profile predicted later academic performance across Key Stages 1-4 (i.e. from 6-7 years of age until 15-16 years of age). There is, however, a dearth of research examining whether the school readiness measure (i.e. the good level of development outcome measure; Office for Standards in Education, Children's Services and Skills, 2014; Pettinger et al., 2019; Public Health England, 2019), can predict later academic outcomes. Evidence that the good level of development measure can predict later educational performance would have important implications for policymakers and researchers in England. More broadly, it would allow the generalization of previous work using the Early Development Index, indicating that holistic school readiness evaluations (assessing both academic and non-academic abilities) are important tools that can be used to predict academic achievement.

An important and novel question that has not yet been examined in the literature is whether holistic school readiness evaluations are equally predictive of later academic achievement in children with and without special educational needs (SEN). If school readiness evaluations are less predictive of academic achievement in children with SEN, this might suggest that such assessments are less useful in this group. In contrast, if the school readiness measure is particularly predictive of later academic achievement in individuals with SEN, this would demonstrate that entering formal education ‘school ready’ is especially important for this group.

Finally, given that children with SEN often have difficulties academically, or with language, communication, social, or motor skills (Department for Health, 2013; Department for Education & Department of Health, 2015; National Center for Education Statistics, 2020), it is plausible that holistic school readiness measures could act as a screening tools to identify children particularly at-risk of SEN. Further strengthening this possibility, previous work

conducted across several countries has observed that children with SEN score more poorly on school readiness evaluations (Janus, 2011). Evidence that holistic school readiness evaluations can be used to identify at risk children would be promising, as there is an urgent need to identify SEN at an earlier timepoint (Hughes et al., 2018). This is challenging, as the term 'SEN' covers a wide range of difficulties (e.g. behavioral, social, communication, physical) that can vary from relatively mild to extremely severe (Lloyd et al., 2009). To date, only a limited number of studies have investigated whether holistic screening tools administered at school entry can identify children at risk of SEN. Several of these studies have used the Early Years Foundation Stage Profile (Snowling, 2003; Norbury et al., 2015; Wright et al., 2019). For instance, Wright et al. (2019) demonstrated that total score could identify individuals at risk of Autism Spectrum Disorder, whilst Norbury et al. (2015) found an association between the good level of development outcome and teacher-ratings of language and behavioral problems. It would, however, be beneficial to examine whether school readiness measures can identify children with SEN more generally (as opposed to specific developmental disorders or difficulties). To the best of our knowledge, only one study to date has examined this thus far (Hughes et al., 2018). In this study, it was found that a teacher-rated evaluation of school readiness (the Brief Early Skills and Support Index), which assesses a variety of abilities at school entry such as 'daily living skills' and 'language and cognition' could correctly classify whether children had SEN. Remarkably, only 3% of children identified as *not* at risk actually had SEN. The evaluation did, however, considerably overidentify children, with only 25% of children classified as 'at-risk', having SEN. As a myriad of factors other than SEN status contribute to school readiness (Nelson et al., 2016), this overidentification is not completely surprising, though it would be useful for research to further examine whether holistic school readiness measures can be used as screening tools to identify individuals at risk of SEN.

The current study examined these issues using the Born in Bradford longitudinal birth cohort study, which is following the lives of over 13,500 children born in Bradford (West Yorkshire, UK) between 2007 and 2011 (Wright et al., 2013). We examined whether the school readiness measure (i.e. the good level of development outcome) derived from the Early Years Foundation Stage Profile could predict scholastic performance at the end of Key Stage 1 (aged 6-7 years of age) in four key subject areas (reading, writing, maths, and science). We also investigated whether the extent to which the good level of development measure predicted later academic performance differed in children with and without SEN. Finally, we examined whether the good level of development outcome could be used as a screening tool to identify children at increased risk of requiring SEN support. The outcomes of this work will answer three novel research questions. First, this study examines whether the school readiness outcome of the Early Years Foundation Stage Profile can predict later academic achievement. This will provide further insights into whether holistic school readiness evaluations can predict academic achievement (Brinkman et al., 2013; Forget-Dubois et al., 2007). Second, this study will investigate the extent to which the predictive validity of school readiness differs as a function of SEN. Finally, the current study will provide further insights into the extent to which holistic assessments at school entry can identify individuals at heightened risk of requiring SEN support.

Based on several perspectives in the literature (Diamond, 2010) and previous research examining school readiness holistically (Brinkman et al., 2013; Forget-Dubois et al., 2007; Hair et al., 2006; Pan et al., 2019), it was predicted that the multidimensional school readiness measure would significantly predict later academic achievement, although we did not have an *a priori* prediction as to whether this would differ in children with and without SEN. Finally, given that children with SEN generally have impairments in academic, language, communication, motor, social or emotional skills (Department for Health, 2013;

Department for Education & Department of Health, 2015; National Center for Education Statistics, 2020), it was hypothesized that the good level of development measure would predict SEN status and identify children who may be at increased risk of needing additional support.

2. Method

2.1. Participants

The Born in Bradford longitudinal birth cohort study (Wright et al., 2013) was used, which contains data from 13,858 children born between 2007 and 2011. Figure 1 displays the the exclusion criteria and the number of participants included in each analysis. From the initial 13,858 participants, 3269 participants were excluded as data were not available for the post-2013 Early Years Foundation Stage Profile assessment. Seven additional participants were excluded as the Early Years Foundation Stage Profile assessment had not been completed in the 'Reception' academic year. A further 2452 were excluded as data were not available for the school variable, SEN variable, or the covariates (sex, pre-term status, English as an additional language, academic month of birth, free school meal status during Reception or Key Stage 1, ethnicity). This left 8130 participants.

[Figure 1 about here]

For the analysis examining SEN status as the outcome variable, further exclusions were made. As we were interested in examining the extent to which the good level of development can predict future SEN support, we excluded participants who began receiving

such support during Reception (i.e. before the Early Years Foundation Stage Profile was conducted; N = 170). We also excluded participants who did not have a SEN status recorded for Reception (N = 4221), as it is likely that some of these children received SEN support during Reception. The SEN analysis was therefore conducted on data from 3739 participants.

For the analysis investigating Key Stage 1 outcomes, 2344 participants (out of 8130) were excluded as data were not available for the post-2016 Key Stage 1 assessments and nine participants were removed as the assessments had not been completed in the typical academic year (i.e. Year 2, at the end of Key Stage 1). Linked data for the Early Years Foundation Stage Profile, Key Stage 1 assessments and the covariates was therefore available for 5777 participants. A small minority of children were then removed as they were disapplied from the curriculum, as allowed under the UK's 2002 Education Act (see Department for Education and Skills, 2006 for further details). This resulted in the removal of 10 participants for the reading outcome, 11 participants for the writing outcome, nine participants for the maths outcome, and 11 participants for the science outcome. A further participant was excluded from the science analysis as they were recorded as being 'absent for long periods of time or recently arrived'. The final analysis was therefore run on 5767 participants for reading, 5766 participants for writing, 5768 participants for maths, and 5765 participants for science.

At the time of analysis, children were aged 8-11 years of age and in Years 4-6 (Key Stage 2). Further demographic information for children included in each analysis is presented in Table 1. To allow comparisons with the broader Born in Bradford sample, the demographic information for the entire cohort (N = 13858) is presented in the supplementary material.

[Table 1 about here]

2.2. Outcome variables

The main outcome variables were (i) teacher-reported performance at the end of Key Stage 1 (Year 2) across four key subject areas (reading, writing, maths, and science) and (ii) whether the child began receiving SEN support after Reception (referred to as 'SEN Status' hereafter). For the reading, writing, and maths assessments, each child could be rated as working "below the standard of the interim pre-key stage standards", "at a pre-key stage standard", "towards the expected standard", "at the expected standard" or "at greater depth within the expected standard". However, some of the lowest outcomes had no, or very few, cases in some of the cells. For instance, when considering reading, all children who were assessed as performing "below the standards of the interim pre-key stage standards" (124 individuals) had not reached a good level of development. Thus, in order to avoid near or complete separation, children rated as working "below the standard of the interim pre-key stage standards", "at a pre-key stage standard", or "towards the expected standard" were coded as performing "below expected". Children who were working "at the expected standard" were coded as performing at "expected" levels, whilst those were working "at greater depth within the expected standard" were coded as performing at "above expected" levels. The science assessment was binary, with children either being rated as having "not met the expected standard" (referred to as "below expected" hereafter) or working "at the expected standard" (referred to as "expected" hereafter).

To form the 'SEN Status' outcome variable, data were retrieved from the local authority school records, which provides information on any changes in SEN support received from the start of schooling (Reception) until now (Year 4-6, depending on the child's age). Such SEN support could have been put in place by the school or have been

mandated following a formal clinical assessment of the child. As discussed in the *Participants* section, we excluded children who were recorded as having SEN during Reception or whose SEN status was first reported after Reception. We also conducted analyses investigating whether the good level of development outcome predicted SEN status at any point during schooling (i.e. from the start of Reception onwards). This analysis included all children (including those with no SEN status recorded in Reception, and those who were recorded as receiving SEN support in Reception). The outcomes of this analysis are reported in the supplementary material.

2.3. Predictor variables

The good level of development measure was calculated using outcomes from the Early Years Foundation Stage Profile assessment. This assessment is made based on observations of the child over the Reception academic year. Children were rated as having reached a good level of development if they exhibited skills that were "expected" or "exceeding" expected levels in the early learning goals that assess "communication and language", "physical development", "personal, social and emotional development", "mathematics", and "literacy" (Standards and Testing Agency, 2017).

For the analyses examining interactions between the good level of development and SEN status), exclusions were not made based on when individuals began receiving SEN support. These analyses therefore examined whether the good level of development outcome predicted academic achieivement differently in children who received SEN support at any time (i.e. from the start of Reception onwards) relative to children who never received such support. Critically, this differs from when SEN was used as an outcome variable, which included only children who began receiving SEN support after Reception.

In all analyses, several factors were controlled for which have been shown to affect both SEN status and academic achievement. These were sex (Department for Education,, 2018a; 2019b), pre-term status (Twilhaar, de Kieviet, Aarnoudse-Moens, van Elburg, & Oosterlaan, 2018; Pettinger et al., 2019), month of birth (Bedard & Dhuey, 2006; Pettinger et al., 2019; Wilson, 2000), maternal education (Dickson, Gregg, & Robinson, 2016; Hollomon, Dobbins, & Scott, 1998), ethnicity (Miller-Cotto & Byrnes, 2016), English as an additional language status (Strand, 2016; Department for Education, 2018a), and free school meals status during Reception and KS1 (Department for Education, 2018a).The perinatal factors (sex, pre-term status, and month of birth) were obtained from medical records. Sex was binary (male/female), whilst pre-term status had four levels based on gestational age (Extremely: <28 weeks, Very: 28 weeks to 31 weeks and 6 days, Moderate-late: 32 weeks to 36 weeks and 6 days, and term: 37 weeks or later; World Health Organization, 2018). Month of birth was converted to academic month of birth, which reflects the month of birth relative to the English academic year (September = Month 1, October = Month 2…. July = Month 11, August = Month 12). Maternal education was coded as the highest educational qualification the mother had completed (< 5 General Certificate of Secondary Education (GCSEs) or equivalent, 5 GCSEs or equivalent, A-Level or equivalent, higher than A-Level, Other, Don't know, or Foreign unknown). Where the mother had completed an educational qualification abroad, this was equivalized to the English system where possible. Maternal education was a self-report measure collected as part of a baseline questionnaire completed during pregnancy or soon after birth. Ethnicity was taken from the child's educational records. The ethnicity groups were formed based on the broader ethnicity categories recommended for use in England by the Office for National Statistics (Office for National Statistics, 2016). These are "White", "Mixed/multiple ethnicity groups", "Asian/Asian British", "Black/African/Caribbean/Black British", "Other ethnic group". English as an additional

language (yes/no) and free school meals status (yes/no) was also taken from educational records. In all analyses, school was included as a random intercept.

2.4. Statistical analysis

2.4.1. Predicting SEN and academic performance

2.4.1.1. Academic outcomes. For the reading, writing, and maths outcomes, a series of multilevel multinomial logistic regressions were conducted to investigate whether the good level of development (reached, not reached) outcome predicted performance on the Key Stage 1 assessment (below expected, expected, above expected). For each outcome, Model 1 investigated whether the good level of development measure predicted academic performance in an unadjusted model. Model 2 then controlled for perinatal and socioeconomic/cultural factors (SEN status, sex, pre-term, academic month of birth, EAL status, maternal education level, free school meals status during Reception and Key Stage 1, and ethnicity). School was included as a random intercept in all models.

A further set of models were then conducted to explore whether an interaction emerged between the good level of development measure and SEN status for each subject outcome. Similarly, this was explored in an unadjusted model, followed by a more complex model which controlled for the perinatal and socioeconomic/cultural factors. For all multilevel multinomial logistic regression models, the base outcome category was set as "Expected". The comparison was therefore between "Expected" vs "Below Expected" and "Expected" vs "Above Expected". To aid in interpretation, the odds ratios for the first comparison ("Expected" vs "Below Expected" were inverted (1/odds ratio)). The odds ratios and confidence intervals therefore reflects a comparison between "Below Expected" vs "Expected".

The analysis was the same for the science outcomes, except a series of multi-level binary logistic regression models were conducted as the outcome had only two levels (below expected vs expected). Across models, the patterns of results were broadly similar for all four domains (reading, writing, maths, and science). Therefore, the results of the analyses for the reading outcome are presented in the main body of text, whilst the analyses for the other academic outcomes are presented in the supplementary material.

2.4.1.2. SEN status. A pair of multi-level binary logistic regressions was conducted to investigate whether the good level of development measure (reached, not reached) predicted SEN status (yes/no). The first model explored whether the good level of development measure predicted SEN status in an unadjusted model, whilst the second model investigated this relation whilst controlling for perinatal and socioeconomic factors (sex, pre-term, academic month of birth, EAL status, maternal education level, free school meals status during Reception and Key Stage 1, and ethnicity). Both models included school as a random intercept.

2.4.2. Classification rates

To further assess the screening utility of the good level of development measure, sensitivity, specificity, positive predictive value, negative predictive value, and correct classification rates were calculated for each outcome variable. Definitions of the terms in relation to each outcome variable are presented in Table 2.

[Table 2 about here]

2.4.3. Handling missing data

Although multiple imputation is commonly used to account for missing data, few statistical packages are currently able to account for data with a multi-level structure. Attempts to impute data whilst retaining the multilevel structure (e.g. using *mice* in R (van Buuren & Groothuis-Oudshoorn 2011), the *mi* function in STATA) were unsuccessful due to poor model convergence. A dummy indicator approach, whereby the multilevel structure is taken into account by including a dummy variable for each cluster (Lüdtke et al., 2017) was not feasible due to the presence of a large number of schools (>170; Wijesuriya et al., 2020) . Two futher approaches were therefore considered: listwise deletion and multiple imputation ignoring the multilevel structure of the data. Although both methods can introduce bias, several researchers have recently warned about the dangers of ignoring the multilevel structure when imputing data (Enders et al., 2015; Grund et al., 2018; Mistler, 2013). Additionally, it has been suggested that imputing multilevel data using a single-level structure may be more hazardous than listwise deletion (van Buuren, 2018). For this reason, we report the results using listwise deletion. Nevertheless, as the best way to deal with missing data is still a topic of debate, some readers may favor multiple imputation using a single-level structure over listwise deletion. We therefore analysed the data a second time after conducting multiple imputation.

Listwise deletion was conducted an on outcome-by-outcome basis, ensuring that participants were only removed from the analysis if they had missing values for variables used in models assessing that outcome. For instance, in the SEN analysis, participants were not excluded if data was missing for the Key Stage 1 academic assessments. Analyses were conducted in STATA. As discussed above, the outcomes of this analysis are reported in text and the supplementary material.

The multiple imputation procedure was conducted using Multivariate Imputations by Chained Equations (MICE). This was implemented in R using the *mice* package (van Buuren

& Groothuis-Oudshoorn 2011). Sex and academic month of birth had no missing values. Binary variables (i.e. good level of development, SEN, free school meals, English as an additional language) were imputed using the "logreg" function, whilst the Key Stage 1 assessments, maternal education, pre-term status, and ethnicity was imputed using the "polyreg" function. The imputation process included all variables except from "school". Five sets of imputed data were produced, with 50 iterations each. The analyses were conducted on each model separately, with results then pooled.

For all analyses, the significance of the main predictor (i.e. the good level of development predictor, and the interaction between good level of development and SEN in some models) did not differ between these different methods of handling missing data. As such, the outcomes following multiple imputation are not reported further. The only exception to this was the interaction between the good level of development and SEN on the Key Stage 1 Science outcome, which was significant using listwise deletion and only approaching significance following multiple imputation ($p = .097$ in the model containing the interaction only, and $p = .082$ controlling for covariates). Caution should therefore be taken when interpreting this outcome, although readers should consider that imputing multilevel data using a single-level structure may produce more bias than listwise deletion (van Buuren, 2018).

3. Results

3.1. Does the school readiness measure predict later reading abilities?

The frequencies and percentages of children achieving each Key Stage 1 reading outcome, depending on whether they reached the good level of development is displayed in Table 3.

[Table 3 about here]

The outcomes for the multilevel multinomial logistic regressions predicting Key Stage 1 reading are presented in Figure 2. The findings are reported firstly for the below expected vs expected comparison, and then for the expected vs above expected comparison.

[Figure 2 about here]

Below expected vs expected. The odds of children who reached a good level of development achieving expected levels on the later reading assessment were significantly higher than for children who did not reach a good level of development. This was observed in both the unadjusted model (Model 1; *Odds Ratio (OR)*: 14.29, *95% Confidence Interval (CI)*: 11.11-16.67, $p < .001$), and in the model controlling for covariates (Model 2; *OR*: 8.33; *95% CI*: 7.14-10.00; $p <. 001$). In Model 2, several of the other predictors were also significant. The odds of children receiving SEN support performing at expected levels were lower than for children who had never received such support (*OR*: 0.23; *95% CI*: 0.19-0.27; $p <. 001$). Relative to children born in September, those born in June had a lower odds of performing at expected levels (*OR*: 0.63; *95% CI*: 0.43-0.94; $p = .024$). Finally, children whose mother had 5 GCSEs (*OR*: 1.25; *95% CI*: 1.02-1.54; $p = .035$), A-Levels (*OR*: 1.45; *95% CI*: 1.11-1.89; $p = .005$), or higher than A-Level (*OR*: 1.85; *95% CI*: 1.45-2.38; $p <. 001$) had higher odds of performing at expected levels relative to children whose mother had the lowest education level (<5 GCSEs).

Expected vs above expected. Children who reached a good level of development had higher odds of performing at above expected levels on the reading assessment than children

who did not reach a good level of development. This was observed in the basic model (*OR*: 8.90; *95% CI*: 6.59-12.02; *p* <. 001), and the model adjusting for covariates (*OR*: 7.08; *95% CI*: 5.20-9.66; *p* <. 001). Several of the covariates were also significant predictors. The odds of achieving above expected levels were lower in children who had SEN (*OR*: 0.56; *95% CI*: 0.40-0.78; *p* = .001), spoke English as an additional language (*OR*: 0.77; *95% CI*: 0.63-0.95; *p* = .013) or received free school meals (*OR*: 0.78; *95% CI*: 0.64-0.95; *p* = .015). Relative to children born in September, those born in all other months had significantly lower odds of performing at above expected levels (except from December; see Figure 3). The odds of achieving above expected levels were higher in females (*OR*: 1.20; *95% CI*: 1.04-1.39; *p* = .011). Relative to children whose mother had fewer than 5 GCSEs, children whose mother had A-Levels (*OR*: 1.30; *95% CI*: 1.01-1.67; *p* = .038), higher than A-Levels (*OR*: 2.34; *95% CI*: 1.87-2.93; *p* <. 001), or an "Other" educational qualification (*OR*: 1.45; *95% CI*: 1.01-2.08; *p* = .044) had a higher odds of performing at above expected levels. Finally, children born extremely premature (*OR*: 9.39; *95% CI*: 1.32-67.02; *p* = .025) had higher odds of performing at above expected levels, although this outcome should be interpreted with extreme caution given the large confidence intervals and the small number of participants with this prematurity outcome.

3.2. Does the effect of school readiness on later reading abilities differ depending on special educational needs status?

Analysis was then completed to investigate whether the extent to which the good level of development measure predicted Key Stage 1 outcomes differed depending on whether the child had ever received SEN support. The frequency and percentage of children achieving the

Key Stage 1 reading outcome is displayed in Figure 3, as a function of good level of development and SEN status.

[Figure 3 about here]

Model 1 examined these effects in an unadjusted model, whilst Model 2 controlled for the perinatal and socioeconomic/cultural variables. Evidence of an interaction might suggest that the extent to which the good level of development measure predicted Key Stage 1 reading differed depending on whether the child had ever received SEN support. The outcomes of these models are presented in Figure 4.

[Figure 4 about here]

Below expected vs expected. The good level of development measure was a significant predictor of later reading performance, with children who reached a good level of development having a higher odds of performing at expected levels (*OR*: 7.14; *95% CI*: 5.88-8.33; $p <$ .001). Children who received SEN support had significantly lower odds of achieving expected reading levels (*OR*: 0.17; *95% CI*: 0.14-0.21; $p <$ .001). There was also a significant interaction between the good level of development measure and SEN status (*OR*: 3.23; *95% CI*: 2.04-5.26; $p <$ .001). These outcomes were similar in Model 2, which controlled for perinatal and socioeconomic/cultural factors (Good level of development *OR*: 6.67; *95% CI*: 5.56-8.33, $p <$ .001; SEN *OR*: 0.17; *95% CI*: 0.14-0.21, $p <$ .001; Interaction *OR*: 3.33; *95% CI*: 2.08-5.56, $p <$ .001). Within Model 2, several of the other variables were also significant predictors. Relative to children whose mother had less than 5 GCSEs, children whose mother had 5 GCSEs (*OR*: 1.23; *95% CI*: 1.01-1.52, $p = .041$), A-Levels (*OR*:

1.45; *95% CI*: 1.11-1.89, $p = .006$), or higher than A-Levels (*OR*: 1.45; *95% CI*: 1.89-2.44, $p <. 001$) had a higher odds of achieving expected reading levels. Children born in June had lower odds of performing at expected reading levels relative to children born in September (*OR*: 0.64; *95% CI*: 0.43-0.95, $p = .028$).

The interaction between the good level of development measure and SEN status is plotted in Figure 5. Bonferroni-corrected pairwise comparisons were conducted to understand this interaction. In Model 1, the odds of achieving expected levels on the Key Stage 1 reading outcome were higher for children who reached a good level of development than children who did not. This was true for children with SEN (*OR* = 23.34, *95% CI* = 12.94-41.68; $p < .001$) and children without SEN (*OR* = 7.10, *95% CI* = 5.53-9.21; $p < .001$), although the odds ratio was larger in children with SEN. For children who did not reach a good level of development, the odds of achieving expected levels on the reading assessment were significantly lower for children with SEN than children with no SEN (*OR* = 0.17, *95% CI* = 0.13-0.23; $p < .001$). For children who reached a good level of development, there was a trend towards lower odds of achieving expected levels in children with SEN relative to children without SEN (*OR* = 0.57, *95% CI* = 0.32-1.01; $p = .057$). Unsurprisingly, children who reached the good level of development and had no SEN had higher odds of achieving expected levels relative to children who did not reach the good level of development and had SEN (*OR* = 40.85, *95% CI* = 30.57-55.15, $p < .001$). Finally, the odds of achieving expected levels on the reading outcome was higher for children with SEN who reached the good level of development than children with no SEN who did not reach a good level of development (*OR* = 4.06, *95% CI*: 2.29-7.17; $p < .001$). These outcomes were similar in Model 2, which controlled for perinatal and socioeconomic/cultural factors (see supplementary material).

[Figure 5 about here]

Expected vs above expected. Reaching a good level of development increased the odds of achieving above expected reading levels (*OR* = 9.26, *95% CI*: 6.62-12.97; $p < .001$). There was no effect of SEN (*OR* = 0.96, *95% CI*: 0.47-1.97; $p = .918$) and the interaction between good level of development and SEN was not significant, although there was a trend (*OR* = 0.45, *95% CI*: 0.20-1.01; $p = .053$). The results were similar in Model 2, which controlled for covariates (Good level of development *OR*: 7.92; *95% CI*: 5.62-11.14, $p <$. 001; SEN *OR*: 0.88; *95% CI*: 0.43-1.82, $p = .738$; Interaction *OR*: 0.50; *95% CI*: 0.22-1.12, $p = .090$). In terms of the other variables, the odds of achieving above expected levels were higher in females than males (*OR* = 1.20, *95% CI*: 1.04-1.38; $p = .012$), and in extremely pre-term children relative to children born at term (*OR* = 8.71, *95% CI*: 1.31-57.86; $p = .025$). As with the primary analysis, this latter finding should be interpreted with caution due to the wide confidence intervals and small number of participants with this pre-term outcome. Relative to children whose mothers had less than 5 GCSEs, the odds of achieving above expected levels were higher in children whose mothers had A-Levels (*OR* = 1.30, *95% CI*: 1.01-1.66; $p = .039$), a qualification higher than A-Levels (*OR* = 2.33, *95% CI*: 1.86-2.92; $p <$ .001), or an "Other" qualification (*OR* = 1.45, *95% CI*: 1.01-2.08; $p = .045$). Having English as an additional language (*OR* = 0.77, *95% CI*: 0.63-0.95; $p = .014$) and receiving free school meals were associated with a lower odds of reaching above expected levels (*OR* = 0.78, *95% CI*: 0.64-0.96; $p = .016$). There was also a clear month of birth effect. Relative to those born in September, children born in other months had a lower odds of achieving above expected levels (except from December; see Figure 5). As the interaction between good level of development and SEN was not significant, post-hoc tests were not conducted.

3.3. Does the school readiness measure predict special educational needs status?

The frequencies and percentages of children that began receiving SEN support after Reception is displayed in Table 3, as a function of the good level of development outcome. Binary logistic regression models were conducted to investigate whether the good level of development measure predicted SEN status. The outcomes are presented in Figure 6. As predicted, children who achieved a good level of development had a significantly lower odds of receiving SEN support in the future. This was observed in both in the unadjusted model (Model 1; *Odds Ratio (OR)*: 0.09, *95% Confidence Interval (CI)*: 0.07-0.12, $p < .001$), and in the model controlling for covariates (Model 2; *OR*: 0.11; *95% CI*: 0.09-0.14; $p <. 001$). Several of the other predictors in Model 2 were also significant. Females had lower odds of receiving SEN support than males (*OR*: 0.55; *95% CI*: 0.45-0.66; $p < .001$), whilst children receiving free school meals had higher odds relative to children who did not (*OR*: 1.27; *95% CI*: 1.02-1.58; $p = .030$). Furthermore, children born moderate-late pre-term (*OR*: 1.55; *95% CI*: 1.05-2.28; $p = .026$) and very pre-term (*OR*: 2.70; *95% CI*: 1.00-7.27; $p = .050$) had higher odds of receiving SEN support than children born at term. Relative to children whose maternal education was equivalent to less than 5 GCSEs, children whose mother's had A-Levels or equivalent (*OR*: 0.71; *95% CI*: 0.51-0.99; $p = .042$), a qualification higher than A-Level (*OR*: 0.69; *95% CI*: 0.51-0.94; $p = .017$), a "Foreign Unknown" qualification (*OR*: 0.37; *95% CI*: 0.14-1.00; $p = .050$), and an "Other" qualification (*OR*: 0.60; *95% CI*: 0.36-0.99; $p = .044$) had lower odds of receiving SEN support. Being Asian/Asian British was also associated with lower odds of receiving SEN support relative to being White (*OR*: 0.72; *95% CI*: 0.51-1.00; $p = .050$).

[Figure 6 about here]

3.4. Can the school readiness measure identify children at risk of poor academic outcomes and special educational needs?

To further assess the utility of the good level of development measure as a screening tool, classification rates were calculated. For the Key Stage 1 reading outcome, sensitivity was 0.83, specificity was 0.79, positive predictive value was 0.56, negative predictive value was 0.94, and correct classification rates were 0.80. For the SEN outcome, sensitivity was 0.83, specificity was 0.66, positive predictive value was 0.36, negative predictive value was 0.94, and correct classification rate was 0.69. As with the main analyses, the classification rates for the writing, maths, and science outcomes are presented in the supplementary material. The classification rates for the SEN analysis including all participants (i.e. regardless of when they began receiving SEN support) is also available in the supplementary material.

4. Discussion

The current study investigated whether a holistic measure of school readiness could predict later academic achievement, and whether this differed as a function of SEN status. The study also explored whether the school readiness outcome could be used to identify children at heightened risk of requiring SEN support. These research questions were assessed using the good level of development outcome from the Early Years Foundation Stage Profile, a statutory assessment conducted universally in England. Reaching a good level of development was associated with higher odds of performing at expected (vs below expected) levels on later educational assessments in key subject areas (reading, writing, maths, and science). This was particularly true for children with SEN. In addition, children who reached a good level of development had increased odds of performing at above expected (vs

expected) levels on the academic assessments. The good level of development measure was also highly predictive of SEN status. To further assess the utility of the school readiness measure, classification rates were calculated. The negative predictive value was high for all outcomes (0.91-0.97), indicating that individuals who reach a good level of development are unlikely to require SEN support or perform at below expected levels on later educational assessments. In contrast, positive predictive value was relatively low for all outcomes (0.36-0.64), demonstrating that around 36-64% of individuals who do not reach a good level of development will not require SEN support or perform poorly academically.

The current study is the first to demonstrate that the school readiness measure of the Early Years Foundation Stage Profile can predict later educational performance. Taken together with studies using the Early Development Index (Brinkman et al., 2013; Forget-Dubois et al., 2017), this indicates that holistic school readiness evaluations can predict educational performance, and that these effects are not limited to a single assessment tool. This finding also suggest that children entering formal education 'school ready' have a substantial advantage over peers who are not 'school ready' (Brinkman et al., 2013; Copple, 1997; Forget-Dubois et al., 2007; Standards and Testing Agency, 2017; Head Start, 2020).

For the first time, the present study examined whether the extent to which holistic school readiness outcomes predict academic achievement differs as a function of SEN status. Indeed, this was consistently observed across all four academic outcomes when comparing the "below expected" to "expected" performance levels, with the association between school readiness and later attainment stronger in children with SEN. Importantly, this indicates that holistic school readiness evaluations are not just useful in children with (or at high risk of) SEN, but that they may actually be more informative in this group. After controlling for covariates, children without SEN who were school ready had approximately 39-91x higher odds of performing at expected (vs below expected) levels on the later academic assessments

than children with SEN who were not school ready. These odds, were however, considerably reduced (to around 1.7-3.2x) when the SEN group passed the school readiness assessment. Thus, although children with SEN are at an academic disadvantage, entering formal education 'school ready' greatly reduces the odds of poor academic outcomes. These findings are in line with previous work, which has suggested that the transition to school is particularly crucial for children with SEN (Janus, 2011).

The current study also demonstrated that outcomes from a holistic school readiness evaluation can predict SEN status. Further analyses of classification rates revealed a high negative predictive value (0.94), indicating that a child who reaches a good level of development is unlikely to require SEN support (6%). The positive predictive value was, however, much lower, with only around 36% of children not reaching a good level of development requiring support for SEN. Such findings are broadly in line with Hughes et al. (2018), who examined whether the Brief Early Skills and Support Index could identify children with SEN. Thus, whilst both evaluations successfully detect most children who will require SEN support, they do appear to overidentify the number of children at risk. Given this overidentification, the question arises as to whether holistic school readiness evaluations could be used practically to identify children who may need SEN support. As both studies evaluated whether the assessments could be used as screening tools, and *not* diagnostic tools, the answer appears to be affirmative. Whilst such assessments are clearly not perfect, they may identify a smaller group of children for further monitoring, which may lead to earlier identification of SEN. This is promising, as there is an urgent need to detect children with these difficulties at an earlier point in order to ensure they have timely access to interventions (Department for Health, 2013; Curran, 2020; Wright et al., 2019). However, given that a very small percentage of children would be missed by both evaluations (around 6% using the good level of development measure, and 3% using the Brief Early Skills and Support Index), it is

clear that educators should not base decisions entirely on school readiness outcomes. Ideally, the results of these evaluations should be combined with regular observations completed by the current teacher, which would allow 'Bayesian' type decision-making regarding the children most at need of additional monitoring or SEN support.

As well as providing theoretical insights into school readiness, the outcomes of this study also have implications for educators and policymakers in England, where the Early Years Foundation Stage Profile is a statutory requirement and conducted universally (Standards and Testing Agency, 2017). Evidence that the good level of development outcome is predictive of later academic achievement, both in children with and without SEN, demonstrates that the evaluation assesses skills that are critical for classroom learning. Such findings are in line with Treadaway (2019), who reported that total score on the evaluation was predictive of educational performance across several key assessment points in childhood and adolescence. The current study is, however, the first to demonstrate that the good level of development outcome, which is often used as the primary measure (e.g. Department for Education, 2017; 2018a; 2019b; Pettinger et al., 2019), can predict later academic attainment. Examination of the classification rates also demonstrated that the evaluation successfully identifies most children who will perform poorly academically (negative predictive value = 0.91-0.97). Educational establishments in England could therefore use this measure to identify children who need close monitoring for early signs of academic difficulties. As these data show, this is particularly important for children that either have, or are at high risk of, SEN. Finally, whilst previous research has shown that the Early Years Foundation Stage Profile can be used to identify children with autism spectrum disorder (Wright et al., 2019), dyslexia (Snowling, 2013) language difficulties (Norbury et al., 2015), or behavioral problems (Norbury et al., 2015), the current study is the first to demonstrate that the

assessment could be used to identify individuals at risk of requiring SEN support more generally.

4.2. Limitations and further research

The current study has several limitations. Firstly, as the design is correlational, causation cannot be established. Further research is therefore needed to examine whether improving school readiness would have downstream effects on educational outcomes. Secondly, although the study controlled for a range of factors (i.e. sex, ethnicity, maternal education, academic month of birth, pre-term status, whether children had English as an additional language, and free school meals status), other variables that were not been controlled for could potentially affect the results. For instance, it is possible that educators assessing children using the Early Years Foundation Stage Profile may have an implicit bias towards children from certain ethnic-minority backgrounds, particularly if they have a different ethnicity themselves (Gilliam et al., 2016). Although the analyses adjusted for the child's ethnicity, teacher demographics were not available and could thus not be controlled for within the current study. It would therefore be useful for further research to investigate this.

A further limitation is that the Key Stage 1 outcome measures used in the current study were teacher-assessments, as opposed to standardized tests of performance. This might therefore limit the conclusions that can be drawn, as these outcomes may be subject to the noise associated with different teachers interpreting the items in different ways. However, recent research has revealed substantial correlations between teacher assessments and standardized test scores at Key Stage 1 (Rimfeld et al., 2019). Nevertheless, it would be useful for research to explore how well the good level of development measure can predict performance on standardized assessments. It would also be beneficial to examine whether the good level of development outcome predicts academic attainment over a longer period of

time, such as at the end of Key Stage 2, when children are 10-11 years of age. Finally, as the current analyses were conducted only on data from Bradford; an area with socioeconomic diversity but a skew towards high levels of deprivation (Wright et al., 2013) and below average levels of academic performance (Department for Education, 2018b), it would be beneficial to replicate the findings from the current study in other populations within England.

4.3. Conclusions

In summary, the current study indicates that a holistic school readiness evaluation that assesses both academic and non-academic abilities can significantly predict later educational performance. A further novel finding was that school readiness was particularly predictive in children with SEN. This suggests that entering formal education school-ready is especially important for this group. Finally, the current study provides evidence that a holistic, school readiness measure can identify children at heightened risk of SEN. This is promising, given the need to identify such children at an earlier time point. Taken together, the current study demonstrates that holistic school readiness evaluations are powerful tools that may allow the earlier identifying of children who are at risk of poor outcomes several years later.

5. References

Ackerman, D. & Barnett W. S. (2005). *Prepared for Kindergarten: What does 'Readiness' Mean?* National Institute for Early Education Research. https://nieer.org/policy-issue/policy-report-prepared-for-kindergarten-what-does-readiness-mean

Aiona, S. (2005). Assessing School Readiness. *Educational Perspectives*, *38*(1), 47-50.

Barbarin, O. A., Early, D., Clifford, R., Bryant, D., Frome, P., Burchinal, M., Howes, C., & Pianta, R. (2008). Parental conceptions of school readiness: Relation to ethnicity, socioeconomic status, and children's skills. *Early Education and Development*, *19*(5), 671-701. https://doi.org/10.1080/10409280802375257

Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, *121*(4), 1437-1472. https://doi.org/10.1093/qje/121.4.1437

Brinkman, S., Gregory, T., Harris, J., Hart, B., Blackmore, S., & Janus, M. (2013). Associations between the early development instrument at age 5, and reading and numeracy skills at ages 8, 10 and 12: a prospective linked data study. *Child Indicators Research*, *6*(4), 695-708. https://doi.org/10.1007/s12187-013-9189-3

van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Copple, C. E (1997). *Getting a good start in school*. National Education Goals Panel, 1997. https://files.eric.ed.gov/fulltext/ED412025.pdf

Claessens, A., & Engel, M. (2013). How Important is Where You Start? Early mathematics knowledge and later school success. *Teachers College Record*, *115*(6), n6.

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*(4), 415-427. https://doi.org/10.1016/j.econedurev.2008.09.003

Curran, H. (2020). *Identifying special educational needs in the early years: perspectives from special educational needs coordinators.* National Association for Special Educational Needs. https://nasen.org.uk/uploads/assets/a610e927-409c-482e-8704847fbf8fd0a8/Identifying-special-educational-needs-in-the-early-years.pdf

Darling-Hammond, L., & Cook-Harvey, C. M. (2018). *Educating the whole child: Improving school climate to support student success.* Learning Policy Institute. https://resourcecentre.savethechildren.net/node/15536/pdf/educating_whole_child_report.pdf

Davies, S., Janus, M., Duku, E., & Gaskin, A. (2016). Using the Early Development Instrument to examine cognitive and non-cognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly*, *35*, 63-75. https://doi.org/10.1016/j.ecresq.2015.10.002

Department for Education & Department of Health. (2015). *Special educational needs and disability code of practice: 0 to 25 years.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/398815/SEND_Code_of_Practice_January_2015.pdf

Department for Education. (2012). *EYFSP attainment by pupil characteristics: 2011 to 2012: National and local authority tables: SFR30/2012.* https://www.gov.uk/government/statistics/eyfsp-attainment-by-pupil-characteristics-in-england-academic-year-2011-to-2012

Department for Education. (2017). *Early years foundation stage profile results in England, 2017.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652602/SFR60_2017_Text.pdf

Department for Education. (2018a). *Early years foundation stage profile results in England, 2018.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/748814/EYFSP_2018_Main_Text.pdf

Department for Education. (2018b). *Bradford Opportunity Area 2017-20. A local plan to improve opportunities for Bradford's children and young people.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696830/Social_Mobility_Delivery_Plan_Bradford_v10_FINAL_WEB.PDF.pdf

Department for Education. (2019a). *Early years foundation stage profile results: 2018 to 2019: EYFSP 2019: tables.* https://www.gov.uk/government/statistics/early-years-foundation-stage-profile-results-2018-to-2019

Department for Education. (2019b). *Early years foundation stage profile results in England, 2019.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/839934/EYFSP_2019_Main_Text_Oct.pdf

Department for Education and Skills (2006). *Disapplication from the National Curriculum (Revised).* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/286565/disapplication_of_the_national_curriculum.pdf

Department of Health (2013). *Annual Report of the Chief Medical Officer 2012: Our Children Deserve Better: Prevention Pays*. www.gov.uk/government/publications/chief-medical-officers-annual-report-2012-our-children-deserve-better-prevention-pays

Diamond, A. (2010). The evidence base for improving school outcomes by addressing the whole child and by addressing skills and attitudes, not just content. *Early Education and Development*, *21*(5), 780-793. https://doi.org/10.1080/10409289.2010.514522

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes? *The Economic Journal*, *126*(596), F184-F231. https://doi.org/10.1111/ecoj.12356

Dockett, S., & Perry, B. (2002). Who's ready for what? Young children starting school. *Contemporary Issues in Early Childhood*, *3*(1), 67-89. https://doi.org/10.2304/ciec.2002.3.1.9

Duncan, G. and Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G.J. Duncan and R.J. Murnane (Eds.), *Whither opportunity: Rising inequality, schools, and children's life chances* (pp.47-70). Russell Sage.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Dunham, S., Lee, E., & Persky, A. M. (2020). The Psychology of Following Instructions and Its Implications. *American Journal of Pharmaceutical Education*, 84(11), ajpe7779. https://doi.org/10.5688/ajpe7779

Enders, C. K., Mistler, S. A., & Keller, B. T. (2016). Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods*, *21*(2), 222. https://doi.org/10.1037/met0000063

Forget-Dubois, N., Lemelin, J. P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development*, *18*(3), 405-426. https://doi.org/10.1080/10409280701610796

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions. *Yale University Child Study Center*, *9*(28). https://medicine.yale.edu/childstudy/zigler/publications/Preschool%20Implicit%20Bias%20Policy%20Brief_final_9_26_276766_5379_v1.pdf

Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010). Fine motor skills and early comprehension of the world: two new school readiness indicators. *Developmental Psychology*, *46*(5), 1008-1017. https://doi.org/10.1037/a0020104

Grund, S., Lüdtke, O., & Robitzsch, A. (2018). Multiple imputation of missing data for multilevel models: Simulations and recommendations. *Organizational Research Methods*, *21*(1), 111-149. https://doi.org/10.1177/1094428117703686

Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., & Calkins, J. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, *21*(4), 431-454. https://doi.org/10.1016/j.ecresq.2006.09.005

Head Start. (2020). *Head Start Approach to School Readiness – Overview.* https://eclkc.ohs.acf.hhs.gov/school-readiness/article/head-start-approach-school-readiness-overview

Heaviside, S., Farris, E., & Carpenter, J. M. (1993). *Public school kindergarten teachers' views on children's readiness for school*. US Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Hollomon, H., Dobbins, D., & Scott, K. (1998). The effects of biological and social risk factors on special education placement: Birth weight and maternal education as an example. *Research in Developmental Disabilities*, *19*(3), 281-294. https://doi.org/10.1016/S0891-4222(98)00002-X

Hughes, C., Foley, S., White, N., & Devine, R. T. (2018). School readiness in children with special educational needs and disabilities: Psychometric findings from a new screening tool, the Brief Early Skills, and Support Index. *British Journal of Educational Psychology*, *88*(4), 606-627. https://doi.org/10.1111/bjep.12206

Isaacs, J., & Magnuson, K. (2011). *Income and education as predictors of children's school readiness.* Center on Children and Families at Brookings. *https://www.brookings.edu/wp-content/uploads/2016/06/1214_school_readiness_isaacs.pdf*

Janus, M. (2011). Impact of impairment on children with special needs at school entry: Comparison of school readiness outcomes in Canada, Australia, and Mexico. *Exceptionality Education International*, *21*(2), 29-44. https://doi.org/10.5206/eei.v21i2.7674

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, *44*(4), 580-589. https://doi.org/10.3758/s13421-015-0579-2

Lin, H.-L., Lawrence, F. R., & Gorrell, J. (2003). Kindergarten teachers' views of children's readiness for school. *Early Childhood Research Quarterly, 18*(2), 225–237. https://doi.org/10.1016/S0885-2006(03)00028-0

Lloyd, J. E., Irwin, L. G., & Hertzman, C. (2009). Kindergarten school readiness and fourth-grade literacy and numeracy outcomes of children with special needs: A population-

based study. *Educational Psychology*, *29*(5), 583-602. https://doi.org/10.1080/01443410903165391

Lopes, L., Santos, R., Pereira, B., & Lopes, V. P. (2013). Associations between gross motor coordination and academic achievement in elementary school children. *Human Movement Science*, *32*(1), 9-20. https://doi.org/10.1016/j.humov.2012.05.005

Lüdtke, O., Robitzsch, A., & Grund, S. (2017). Multiple imputation of missing data in multilevel designs: A comparison of different strategies. *Psychological methods*, *22*(1), 141.

Miller-Cotto, D., & Byrnes, J. P. (2016). Ethnic/racial identity and academic achievement: A meta-analytic review. *Developmental Review*, *41*, 51-70. https://doi.org/10.1016/j.dr.2016.06.003

Mistler, S. A. (2013). A SAS macro for applying multiple imputation to multilevel data. In *Proceedings of the SAS Global Forum* (pp. 1-8). San Francisco, CA: Contributed Paper (Statistics and Data Analysis).

National Center for Education Statistics (2020). Students With Disabilities. https://nces.ed.gov/programs/coe/indicator_cgg.asp

Nelson, B. B., Dudovitz, R. N., Coker, T. R., Barnert, E. S., Biely, C., Li, N., Szilagyi, P. G., Larson, K., Halfon, N., Zimmerman, F. J., & Chung P. J. (2016). Predictors of poor school readiness in children without developmental delay at age 2. *Pediatrics*, *138*(2), e20154477. https://doi.org/10.1542/peds.2015-4477

Norbury, C. F., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2016). Younger children experience lower levels of language competence and academic progress in the first year of school: evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(1), 65-73. https://doi.org/10.1111/jcpp.12431

Office for National Statistics (2016). *Ethnic group, national identity and religion.* https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion

Office for Standards in Education, Children's Services and Skills. (2014). *Are you ready? Good practice in school readiness*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/418819/Are_you_ready_Good_practice_in_school_readiness.pdf

Office for Standards in Education, Children's Services and Skills. (2017). *Bold beginnings: The Reception curriculum in a sample of good and outstanding primary schools.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/663560/28933_Ofsted_-_Early_Years_Curriculum_Report_-_Accessible.pdf

Pagani, L. S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: a French Canadian replication and extension. *Developmental Psychology*, *46*(5), 984-994. https://doi.org/10.1037/a0018881

Pan, Q., Trang, K. T., Love, H. R., Templin, J. (2019). School Readiness Profiles and Growth in Academic Achievement. *Frontiers in Education, 4*, 1-17. http://dx.doi.org/10.3389/feduc.2019.00127

Parker, J. D., Creque Sr, R. E., Barnhart, D. L., Harris, J. I., Majeski, S. A., Wood, L. M., Bond, B. J., & Hogan, M. J. (2004). Academic achievement in high school: does emotional intelligence matter? *Personality and Individual Differences*, *37*(7), 1321-1330. https://doi.org/10.1016/j.paid.2004.01.002

Pettinger, K. J., Kelly, B., Sheldon, T. A., Mon-Williams, M., Wright, J., & Hill, L. J. B. (2019). Starting school: educational development as a function of age of entry and

prematurity. *Archives of Disease in Childhood*, http://dx.doi.org/10.1136/archdischild-2019-317124

Public Health England (2019). *Public Health Profiles: Indicator Definition and Supporting Information*. https://fingertips.phe.org.uk/search/school%20readiness#page/6/gid/1/pat/6/par/E12000004/ati/102/are/E06000015/iid/90631/age/34/sex/4

Rimfeld, K., Malanchini, M., Hannigan, L. J., Dale, P. S., Allen, R., Hart, S. A., & Plomin, R. (2019). Teacher assessments during compulsory education are as reliable, stable and heritable as standardized test scores. *Journal of Child Psychology and Psychiatry. 60*(12), 1278-1288. https://doi.org/10.1111/jcpp.13070

Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *The Journal of Early Adolescence*, *14*(2), 226-249. https://doi.org/10.1177/027243169401400207

Snow, K. L. (2006). Measuring school readiness: Conceptual and practical considerations. *Early Education and Development*, *17*(1), 7-41. https://doi.org/10.1207/s15566935eed1701_2

Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *Journal of Research in Special Educational Needs*, *13*(1), 7-14. https://doi.org/10.1111/j.1471-3802.2012.01262.x

Standards and Testing Agency. (2017). *Early years foundation stage profile 2018 handbook.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/669079/Early_years_foundation_stage_profile_2018_handbook.pdf

Stormont, M., Herman, K. C., Reinke, W. M., King, K. R., & Owens, S. (2015). The Kindergarten Academic and Behavior Readiness Screener: The utility of single-item teacher ratings of kindergarten readiness. *School Psychology Quarterly*, *30*(2), 212-228. https://doi.org/10.1037/spq0000089

Strand. (2016). *Educational Outcomes Among Children with English as an Additional Language (EAL).* http://migrationobservatory.ox.ac.uk/wp-content/uploads/2016/06/Briefing-Educational_Outcomes_EAL-1.pdf

Twilhaar, E. S., de Kieviet, J. F., Aarnoudse-Moens, C. S., van Elburg, R. M., & Oosterlaan, J. (2018). Academic performance of children born preterm: a meta-analysis and meta-regression. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *103*(4), F322-F330. http://dx.doi.org/10.1136/archdischild-2017-312916

Treadaway, M. (2019). *How attainment gaps emerge from Foundation Stage to Key Stage 4, part one*. https://ffteducationdatalab.org.uk/2019/10/how-attainment-gaps-emerge-from-foundation-stage-to-key-stage-4-part-one/

UK Government (N.D.a). *School admissions.* https://www.gov.uk/schools-admissions/school-starting-age

UK Government (N.D.b). *Early years foundation stage.* https://www.gov.uk/early-years-foundation-stage

Waterman, A. H., Atkinson, A. L., Aslam, S. S., Holmes, J., Jaroslawska, A., & Allen, R. J. (2017). Do actions speak louder than words? Examining children's ability to follow instructions. *Memory & Cognition*, *45*(6), 877-890. https://doi.org/10.3758/s13421-017-0702-7

Wijesuriya, R., Moreno-Betancur, M., Carlin, J. B., & Lee, K. J. (2020). Evaluation of approaches for multiple imputation of three-level data. *BMC medical research methodology*, *20*(1), 1-15. https://doi.org/10.1186/s12874-020-01079-8

Wilson, G. (2000). The effects of season of birth, sex and cognitive abilities on the assessment of special educational needs. *Educational Psychology*, *20*(2), 153-166. https://doi.org/10.1080/713663714

World Health Organization. (2018). *Preterm birth.* https://www.who.int/news-room/fact-sheets/detail/preterm-birth

Wright, B., Mon-Williams, M., Kelly, B., Williams, S., Sims, D., Mushtaq, F., Sohal, K., Blackwell, J, E., & Wright, J. (2019). Investigating the association between early years foundation stage profile scores and subsequent diagnosis of an autism spectrum disorder: a retrospective study of linked healthcare and education data. *BMJ Paediatrics Open*, *3*(1). 10.1136/bmjpo-2019-000483

Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N., Fairley, L., Lawlor, D. A., Parslow, R., Petherick, E., Pickett, K., E., Waiblinger, D., & West, J. (2013). Cohort profile: the Born in Bradford multi-ethnic family cohort study. *International Journal of Epidemiology*, *42*(4), 978-991. https://doi.org/10.1093/ije/dys112

Ziol-Guest, K. M., & McKenna, C. C. (2014). Early childhood housing instability and school readiness. *Child Development*, *85*(1), 103-113. https://doi.org/10.1111/cdev.12105

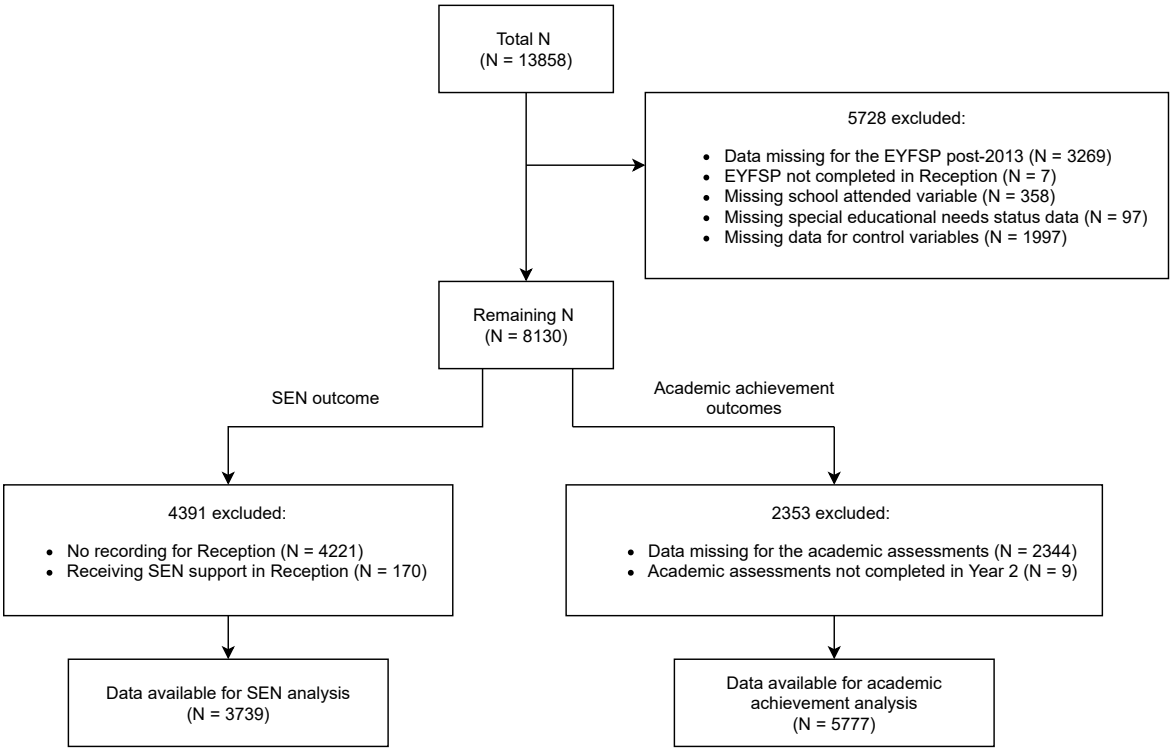

Figure 1. The exclusion criteria used and participant numbers used within the analyses. Note that several further participants were removed from the Key Stage 1 assessment analysis (see participant section for more details).

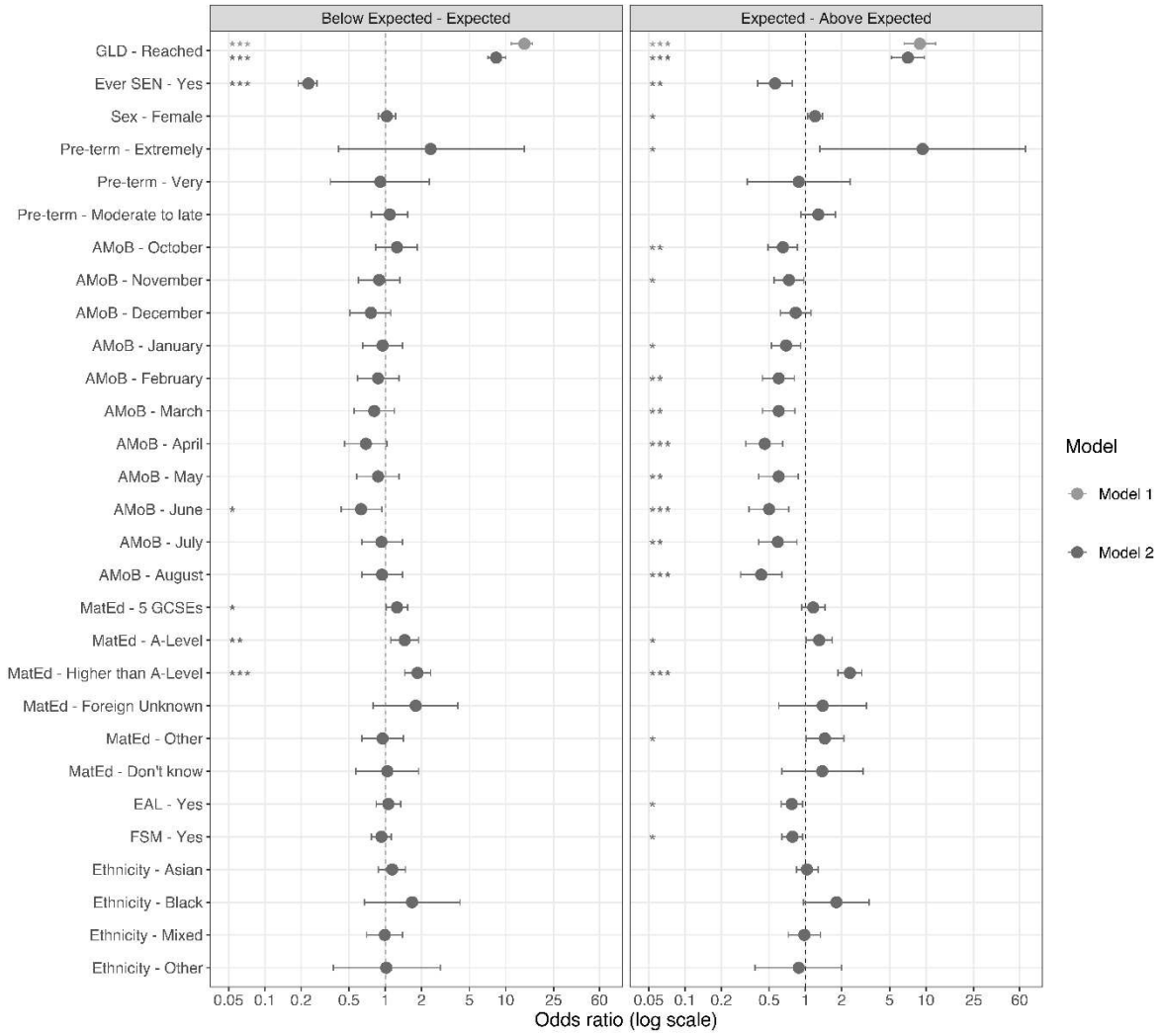

Figure 2. The odds ratios, 95% confidence intervals and p-values for the multilevel multinomial logistic regressions predicting Key Stage 1 reading outcome (below expected/expected/above Expected). Model 1 examined whether good level of development predicted Key Stage 1 reading in an unadjusted model, whilst Model 2 adjusted for covariates. Circular points reflect odds ratios, the horizontal bars display the 95% confidence intervals, and the stars denote statistical significance (*** = $p < .001$, ** = $p <. 01$, * = $p < .05$).

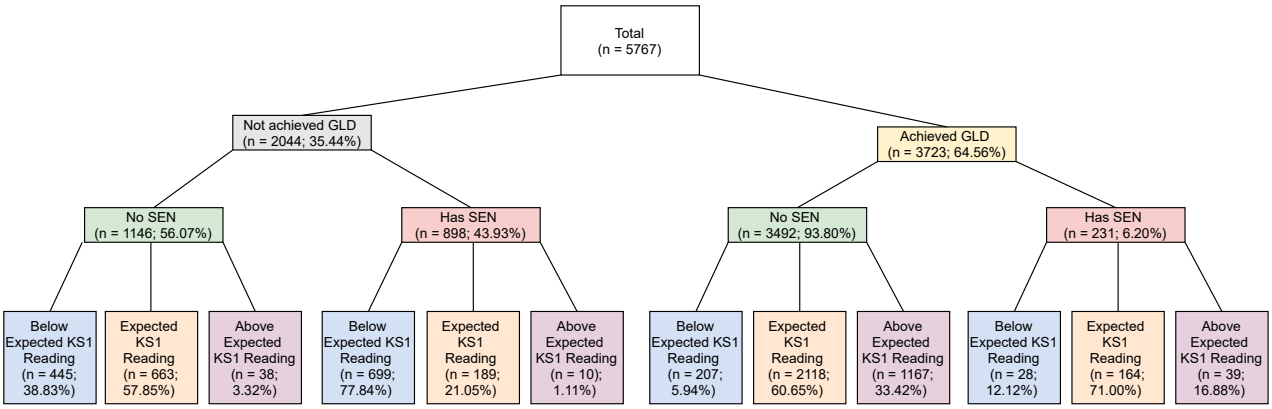

Figure 3. The frequency and percentage of children achieving each Key Stage 1 reading outcome, as a function of good level of development and SEN status.

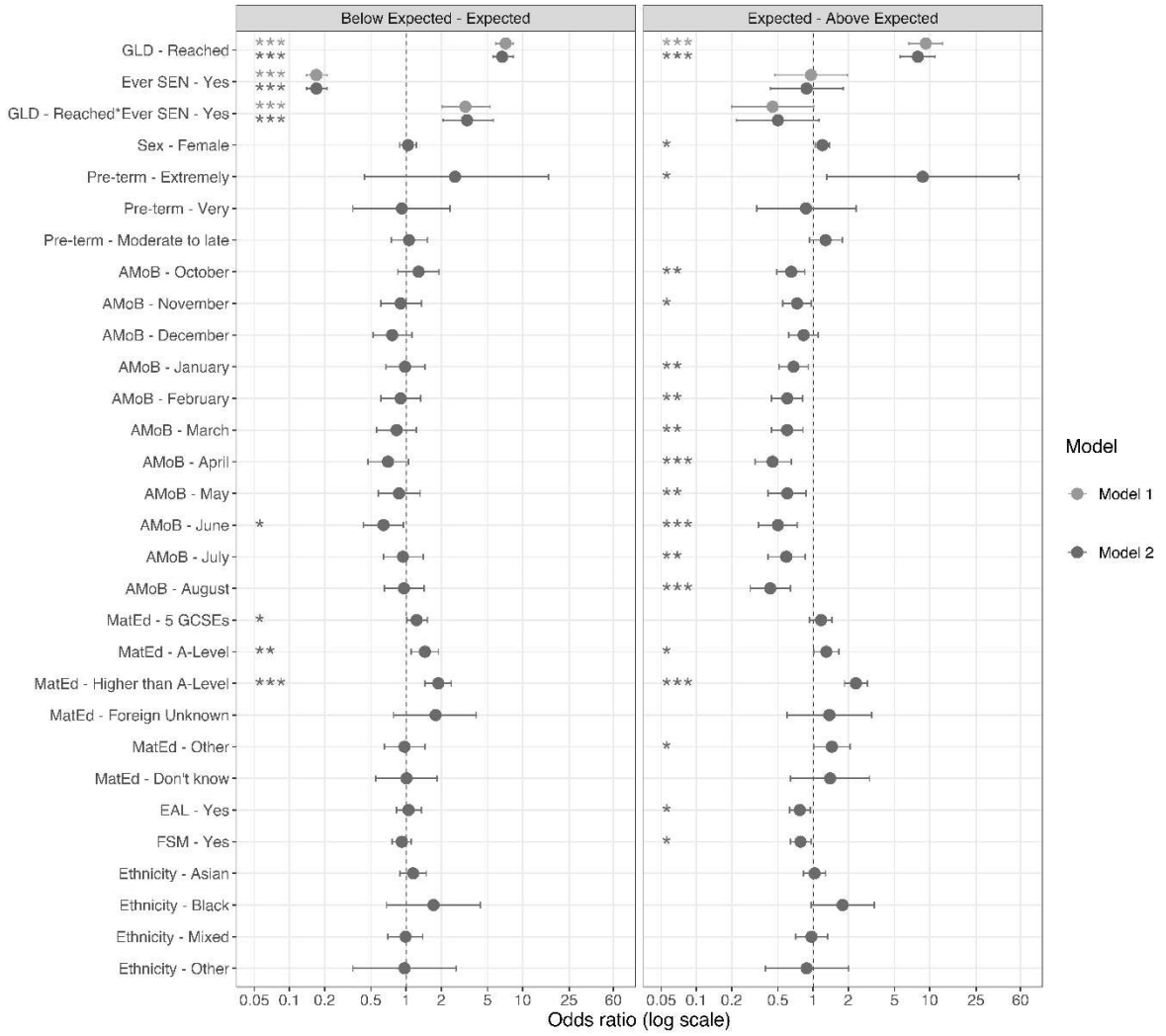

Figure 4. The odds ratios, 95% confidence intervals and p-values for the multilevel multinomial logistic regression models predicting Key Stage 1 reading (Below Expected/Expected/Above Expected). Model 1 examined whether good level of development and SEN status predicted Key Stage 1 reading abilities, and whether an interaction emerged between them. Model 2 then adjusted for the perinatal and socioeconomic covariates. Circular points reflect the odds ratio, horizontal bars display 95% confidence intervals, and stars denote statistical significance (*** = $p < .001$, ** = $p <. 01$, * = $p < .05$).

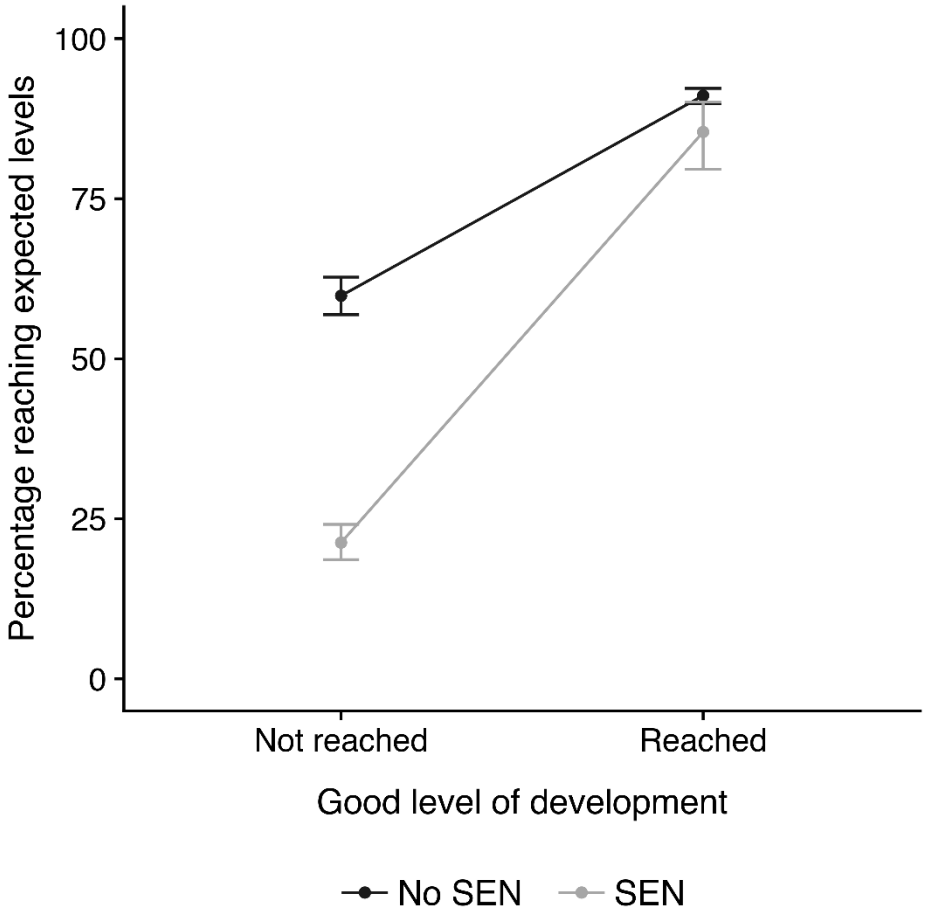

Figure 5. The percentage of children who reached expected levels on the Key Stage 1 reading assessment as a function of the good level of development outcome and SEN status. Percentages were calculated by dividing the number of children who performed at expected levels on the Key Stage 1 reading assessment by the number of children who performed at below expected or expected levels. Error bars reflect binomial 95% confidence intervals, calculated using the Clopper-Pearson exact method. The upper and lower confidence intervals were multiplied by 100, such that values reflect percentages rather than proportions.

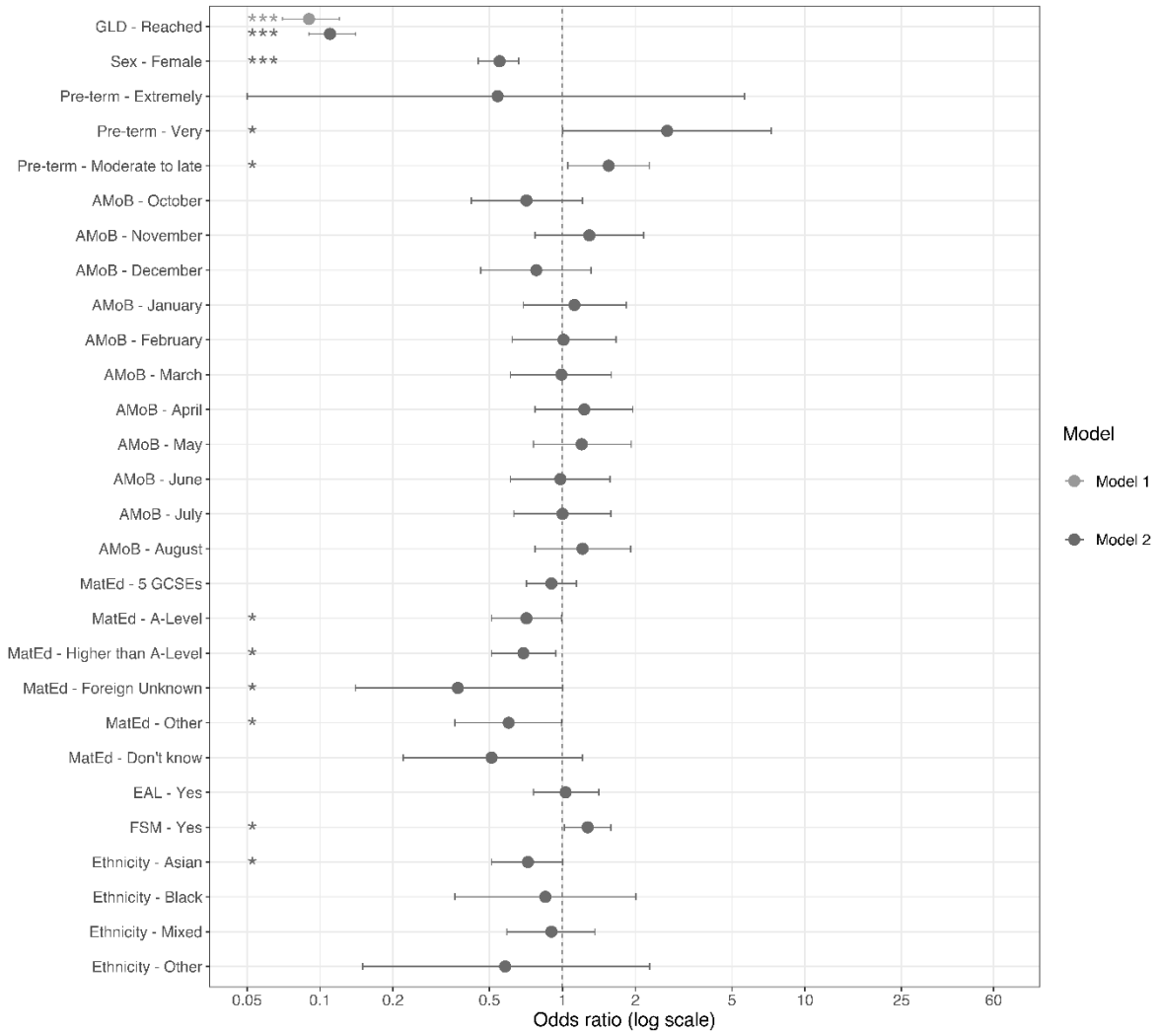

Figure 6. The odds ratios, 95% confidence intervals and p-values for the multilevel binary logistic regression predicting SEN status (yes/no). Model 1 examined whether the good level of development measure predicted SEN status in an unadjusted model, whilst Model 2 then adjusted for perinatal and socioeconomic/cultural covariates. The circular points reflect odds ratios, whilst the horizontal bars denote the 95% confidence intervals. The stars reflect statistical significance (*** = $p < .001$, ** = $p <. 01$, * = $p < .05$).

Table 1

Demographic information for the participants included in the analysis predicting SEN status (column 2) and performance on the Key Stage 1 reading, writing, maths, and science assessments (column 3-6). Percentages are displayed in brackets. These may not add to 100% due to rounding.

	SEN (N = 3739)	Reading (N = 5767)	Writing (N = 5766)	Maths (N = 5768)	Science (N = 5765)
Good Level of Development - Yes	2122 (56.75%)	3723 (64.56%)	3723 (64.57%)	3723 (64.55%)	3722 (64.56%)
SEN status - Yes	704 (18.83%)	1129 (19.58%)	1128 (19.56%)	1130 (19.59%)	1128 (19.57%)
Key Stage 1 outcome					
Below	-	1379 (23.91%)	1647 (28.56%)	1368 (23.72%)	1051 (18.23%)
Expected	-	3134 (54.34%)	3265 (56.63%)	3297 (57.16%)	4714 (81.77%)
Above	-	1254 (21.74%)	854 (14.81%)	1103 (19.12%)	-
Gender - Female	1975 (52.82%)	2837 (49.19%)	2837 (49.2%)	2837 (49.19%)	2837 (49.21%)
Academic month of birth					
September	348 (9.31%)	583 (10.11%)	583 (10.11%)	583 (10.11%)	583 (10.11%)
October	309 (8.26%)	593 (10.28%)	593 (10.28%)	593 (10.28%)	592 (10.27%)
November	267 (7.14%)	555 (9.62%)	555 (9.63%)	555 (9.62%)	554 (9.61%)
December	290 (7.76%)	545 (9.45%)	545 (9.45%)	545 (9.45%)	544 (9.44%)
January	288 (7.70%)	592 (10.27%)	593 (10.28%)	593 (10.28%)	593 (10.29%)
February	271 (7.25%)	491 (8.51%)	491 (8.52%)	491 (8.51%)	492 (8.53%)
March	322 (8.61%)	488 (8.46%)	488 (8.46%)	488 (8.46%)	487 (8.45%)
April	358 (9.57%)	391 (6.78%)	391 (6.78%)	391 (6.78%)	391 (6.78%)
May	319 (8.53%)	355 (6.16%)	355 (6.16%)	355 (6.15%)	355 (6.16%)
June	318 (8.50%)	374 (6.49%)	374 (6.49%)	374 (6.48%)	374 (6.49%)
July	319 (8.53%)	407 (7.06%)	406 (7.04%)	407 (7.06%)	407 (7.06%)
August	330 (8.83%)	393 (6.81%)	392 (6.8%)	393 (6.81%)	393 (6.82%)
English as an additional language - Yes	1636 (43.76%)	2498 (43.32%)	2496 (43.29%)	2498 (43.31%)	2497 (43.31%)
Pre-term*					
Extremely	5 (0.13%)	9 (0.16%)	9 (0.16%)	9 (0.16%)	9 (0.16%)
Very	22 (0.59%)	35 (0.61%)	35 (0.61%)	35 (0.61%)	35 (0.61%)
Moderate-Late	179 (4.79%)	291 (5.05%)	291 (5.05%)	291 (5.05%)	290 (5.03%)
Term	3533 (94.49%)	5432 (94.19%)	5431 (94.19%)	5433 (94.19%)	5431 (94.21%)
Maternal education†					
<5 GCSEs	893 (23.88%)	1252 (21.71%)	1252 (21.71%)	1253 (21.72%)	1251 (21.7%)
5 GCSEs	1207 (32.28%)	1833 (31.78%)	1832 (31.77%)	1833 (31.78%)	1834 (31.81%)
A-Level	488 (13.05%)	935 (16.21%)	935 (16.22%)	935 (16.21%)	933 (16.18%)
>A-Level	853 (22.81%)	1348 (23.37%)	1348 (23.38%)	1348 (23.37%)	1348 (23.38%)
Foreign unknown	39 (1.04%)	52 (0.9%)	52 (0.9%)	52 (0.9%)	52 (0.9%)

Other	216 (5.78%)	273 (4.73%)	273 (4.73%)	273 (4.73%)	273 (4.74%)
Don't know	43 (1.15%)	74 (1.28%)	74 (1.28%)	74 (1.28%)	74 (1.28%)
Free school meals - Yes	942 (25.19%)	1212 (21.02%)	1212 (21.02%)	1213 (21.03%)	1212 (21.02%)
Ethnicity					
White	1499 (40.09%)	2251 (39.03%)	2252 (39.06%)	2252 (39.04%)	2251 (39.05%)
Asian	1961 (52.45%)	3056 (52.99%)	3054 (52.97%)	3056 (52.98%)	3054 (52.97%)
Black	48 (1.28%)	62 (1.08%)	62 (1.08%)	62 (1.07%)	62 (1.08%)
Mixed	207 (5.54%)	356 (6.17%)	356 (6.17%)	356 (6.17%)	356 (6.18%)
Other	24 (0.64%)	42 (0.73%)	42 (0.73%)	42 (0.73%)	42 (0.73%)

SEN = special educational needs, GCSE = General Certificate in Secondary Education

- Not included as a factor in the analysis

* defined by the World Health Organization (2018). Extremely: < 28 weeks gestation, Very: 28 weeks to 31 weeks and 6 days gestation, Moderate-late: 32 weeks to 36 weeks and 6 days gestation, and term: 37 weeks or later

† Equivalized

Table 2

Definitions used for the classification rates analysis for the SEN status and academic outcome variables.

	SEN status	Academic outcomes
Sensitivity	The proportion of children who needed SEN support that did not earlier reach a good level of development.	The proportion of children who experienced "below expected" that did not earlier reach a good level of development.
Specificity	The proportion of children that did not need SEN support who had earlier reached a good level of development.	The proportion of children that exhibited "expected" educational performance who had earlier reached a good level of development.
Positive predictive value	The probability that a child who did not reach a good level of development required SEN support.	The probability that a child who did not reach a good level of development experienced "below expected" Key Stage 1 outcomes.
Negative predictive value	The probability that a child who reached a good level of development did not require SEN support.	The probability that a child who reached a good level of development achieved an "expected" level on the academic outcomes.
Correct classification rates	The proportion of children correctly classified as at risk/not at risk of needing SEN support based on the good level of development.	The proportion of children correctly classified as at risk/not at risk of below expected educational performance based on the good level of development.

Table 3

The number of children who experienced each special educational needs status outcome (yes/no) and, Key Stage 1 reading outcome (below expected, expected, above expected), as a function of the good level of development measure (reached/not reached). For the SEN analysis, all children were recorded as not receiving SEN support in Reception. "No SEN" thus refers to children who never received SEN support, whilst "SEN" refers to children who began receiving SEN support after Reception. The values in brackets reflect the percentage of children experiencing each outcome dependent on whether they reached a good level of development. Percentages may not add to 100% due to rounding.

	Good level of development outcome		
	Not Reached	Reached	*Total*
SEN status			
No SEN	1035 (64.01%)	2000 (94.25%)	*3035*
SEN	582 (35.99%)	122 (5.75%)	*704*
Total	1617	2122	*3739*
Key Stage 1 reading outcome			
Below expected	1144 (55.97%)	235 (6.31%)	*1379*
Expected	852 (41.68%)	2282 (61.29%)	*3134*
Above expected	48 (2.35%)	1206 (32.39%)	*1254*
Total	*2044*	*3723*	*5767*