(Mis)perceptions about children [*]

Anastasiia Suvorova [†]

October 27, 2023

Latest version available here.

Abstract

Policymakers, schools, and parents rely on teachers' assessments of child development to inform decisions about investments in children. I show that teachers' perceptions of children's developmental delays relative to children of the same age are biased and depend systematically on the average development level of other children in the neighbourhood. I quantify the magnitude of the reference bias in teachers' assessments of non-cognitive and cognitive skills using objective measures of non-cognitive skills and language development, evaluated by psychologist-trained interviewers in the Longitudinal Study of Australian Children. I estimate a measurement system of teachers' recognition of children's deficits as a function of children's measured development and neighbourhood average development levels, and I show that teachers in neighbourhoods with lower average levels of non-cognitive development are less likely to report delays in both cognitive and non-cognitive dimensions of child development. Further, maternal perceptions of their children's non-cognitive development are influenced by the information about deficits that teachers convey. Teachers' misperceptions affect investment in remedial services including children's learning and behavioural therapy, and tutoring, as well as parental attitudes toward their children. Finally, I show that teachers' education improves the ability of teachers to identify children with deficits.

[*]I would like to express my gratitude to Audra Bowlus, Lance Lochner, and Sergio Ocampo Diaz, for their guidance and support. I am thankful to Rory McGee, Nirav Mehta, and the participants of the AMM Reading Group for their excellent advice. All errors are my own.

[†]Department of Economics, Western University, London, ON, N6A 5C2, e-mail: asuvorov@uwo.ca

1 Introduction

Correctly identifying developmental deficits is important for schools and families when they make decisions to invest in children's development, from parenting time and reading with the child to hiring professional help like tutors (Kinsler and Pavan, 2021, Dizon-Ross, 2019). Because of this, biased perceptions of child development may lead to suboptimal investment strategies and disrupt healthy child development. In the search for information about children's progress, families and school officials commonly rely on teachers' professional judgements. Teachers' role is particularly important in assessing non-cognitive (non-cognitive) development, where the results of standardized tests are usually not available to complement teachers' reports.

I show that teachers' assessments of child development are biased and depend systematically on the average level of non-cognitive development of other children in the neighbourhood. Teachers of children who live in neighbourhoods with low average levels of non-cognitive development are less likely to indicate deficits in the non-cognitive and cognitive dimensions of development. I show that these misperceptions have a cascading effect on parental perceptions of children and children's home and school environments. Teachers' recognition of developmental deficits affects mothers' perceptions about their children. In turn, these perceptions impact children's uptake of therapy services and family investments in children, including tutoring, parenting styles, and aspirations.

Quantifying the bias in teachers' assessments requires objective measures of child development available for a representative sample of children. For cognitive or academic skills, measures based on standardized tests are widely available and the emerging literature has used these measures to show that teachers' assessments of cognitive skills are biased relative to the average level of cognitive skills in schools (Kinsler and Pavan, 2021, Elder and Zhou, 2021).[3] In contrast to cognitive skills, bias in teachers' assessment of non-cognitive skills and the impact of such bias on children's lives has not been directly quantified from the data due to the lack of objective measures.[4] At the same time, non-cognitive skills have been shown to be of paramount importance

[3]While suffering from important limitations like the strong association with students' effort during testing (Zamarro, Hitt, and Mendez, 2019) or the reinforcement of unequal opportunities (Reeves and Halikias, 2017), test scores provide an objective measure of children's positions in the distribution of cognitive skills for similar-aged children.

[4]Elder and Zhou (2021) aim to quantify the potential reference bias in teachers' reports of non-cognitive skills,

for children's dynamic skill accumulation and life outcomes (Cunha, Heckman, and Schennach, 2010, Deming, 2022).

I overcome this challenge by using direct observations from psychologist-trained interviewers who evaluate children's non-cognitive skills during face-to-face interviews for the Longitudinal Study of Australian Children (LSAC), a nationally representative survey of 10,000 children followed biennially starting in 2004. These observations provide objective measures of children's non-cognitive development in addition to measures of cognitive skills based on language tests that are also available in the dataset. To quantify the reference group bias in teachers' perceptions of non-cognitive and cognitive skills, I use an approach similar to Kinsler and Pavan (2021) and Elder and Zhou (2021) and estimate a measurement system of teachers' deficit recognition allowing it to depend both on children's individual development and on the average level of child development in the neighbourhood.

I use teachers' evaluations of delays in children's cognitive and non-cognitive development relative to other children of a similar age as measures of teachers' perceptions. I show that teachers in neighbourhoods where kids, on average, have lower objective measures of non-cognitive development are less likely to report non-cognitive and cognitive deficits in children conditional on objective measures of children's skills. Specifically, teachers in neighbourhoods at the top quartile of average children's non-cognitive development are actually 1.5 percentage points more likely to report delays in children's non-cognitive development compared to teachers in bottom-quartile neighbourhoods, while they would be up to 10 percentage points less likely to report delays in non-cognitive development if their assessment was free from the reference group bias.

My results also indicate that teachers' assessments of children's cognitive delays depend on the average level of other children's cognitive development in the neighbourhood, consistent with the findings in Kinsler and Pavan (2021) and Elder and Zhou (2021). These findings have important implications for governments aiming to identify disadvantaged areas based on nationwide teacher evaluation statistics like the Australian Early Development Census as the prevalence

however, in the absence of objective measures they rely on restrictive assumptions about the unobserved distribution of non-cognitive skills or the magnitude of the reference bias. By contrast, I take a data-driven approach to quantifying the reference bias in evaluations of non-cognitive skills.

of child developmental deficits is underestimated in disadvantaged and overestimated in advantaged areas. Similarly, researchers relying on assessments of children performed by teachers and parents need to account for the reference bias in their interpretations of estimated gaps in skills across children and changes in skills with age.

I also explore the role of teacher qualifications and classroom characteristics in the identification of developmental deficits in the spirit of studies analyzing the impact of educational program characteristics on student outcomes.[5] I find that teachers' judgement about children's developmental delays does not improve significantly with teaching experience or smaller class sizes. However, more educated teachers have a higher probability of indicating developmental deficits in children with low objective measures of development. For the subsample of children with low objective measures of non-cognitive development, teachers with a university degree are 6.2 percentage points more likely to indicate delays in non-cognitive development compared to teachers with certificates or diplomas. Similarly, for children with low objective measures of cognitive skills, teachers with university training are 5.1 percentage points more likely to report delays. This advantage in deficit recognition is driven by the stronger association between children's objective cognitive abilities measure and deficit recognition in teachers with university training. This implies that the training can be an effective tool assisting teachers in deficit recognition.

I show that mothers update their perceptions when teachers inform them about non-cognitive deficits in children. While the role of schools in mothers' perceptions about children's academic progress has been studied in the literature (Dizon-Ross, 2019, Doss, Fahle, Loeb, and York, 2019), the relationship between teachers' and mothers' perceptions about non-cognitive skills is less understood. I use information contained in the LSAC about mothers being contacted by schools about children's behavioural problems to quantify the effect of teachers' deficit identification on mothers. Being contacted by the school increases the probability that mothers perceive the child to have non-cognitive delays by almost 12 percentage points.

Moreover, the recognition of children's non-cognitive delays by mothers is related to parental

[5]For example, Chetty, Friedman, Hilger, Saez, Schanzenbach, and Yagan (2011) document that students who had a more experienced teacher in kindergarten have higher earnings. Goldhaber and Brewer (2000) show that teachers' certification has a significant effect on students' test scores.

attitudes towards children and family investment decisions. To investigate the impact of mothers' perceptions on parental choices, I adjust the reduced-form approach used in Kinsler and Pavan (2021) to investigate the role of mothers' perceptions about non-cognitive skills for investments, parenting practices, and attitudes that were shown to be productive for non-cognitive development (Fiorini and Keane, 2014, Falk, Kosse, Pinger, Schildberg-Hörisch, and Deckers, 2021). I estimate value-added regressions of family investment choices on maternal deficit recognition accounting for the persistence in maternal perceptions and investment. I show that mothers reporting non-cognitive delays in children differ in their parental attitudes and expectations about children, with mothers who recognize children's deficits engaging in more angry and less warm parenting and having lower educational aspirations for their children. By contrast, deficit recognition by mothers is associated with higher use of tutoring.

Additionally, I show that the identification of developmental delays in children by teachers and mothers determines whether children take advantage of therapy directed at non-cognitive or cognitive skills. In my data, teachers and mothers report whether children use school or community services like behavioural therapy or psychological evaluation, as well as speech and learning therapy. My findings imply that children whose teachers identify non-cognitive delays are 7 percentage points more likely to use behavioural therapy or undergo a psychological evaluation. They are also 7 percentage points more likely to use learning or speech therapy.

On one hand, recognizing deficits in children induces parents and teachers to reach out for professional help for their children. Because of this, reference bias in perceptions can lead to the reinforcement of skill gaps between advantaged and disadvantaged areas through gaps in investments in children. On the other hand, deficit recognition is associated with a lower quality of parenting, therefore, overestimation of deficits due to the reference bias may also lead to negative consequences for children who are developmentally on track. In this way, my work sheds light on sources of differences in child environment across parental socioeconomic status (SES) and the role of neighbourhoods. I emphasize the role of neighbourhood-related information frictions in explaining differences in parental behaviour in addition to other factors including resource constraints, preferences, and differences in perceptions about returns to investment that have been shown to drive the gaps in parental investment across the SES and neighbourhoods.[6]

My paper adds to the broader literature on the relationship between parental perceptions and parenting choices. An important strand of this literature has explored the relationship between differences in beliefs about returns to various types of parental investments and actual investment choices.[7] Instead, I focus on the assessment of child development and its role in family investments and parental attitudes. I contribute to research investigating the relationship between parental perceptions of children's skills and their decisions regarding children's environment by exploring perceptions and parenting choices related to non-cognitive skills.[8]

This paper proceeds by first documenting the source of distortions in perceptions about child development and then exploring its consequences. Section 2 describes the data that allows me to investigate the role of reference bias in teachers' perceptions and its effects. Section 3 presents a conceptual framework where the reference group bias in teachers' perceptions can result in a cascading effect on children's environments. Sections 4 and 5 describe the analysis and conclusions for the role of the local environment in teachers' and mothers' perceptions of child development. Sections 6 and 7 proceed by exploring the impact of these perceptions on the children's environment. Section 8 provides some final discussion.

2 Data

The data for this project come from the LSAC, a national study of children in Australia that tracks childhood environments, development, and life course trajectories. The survey commenced in 2004 with participating families interviewed once every two years. It follows the development of two cohorts of children: the baby cohort (B-cohort), which includes 5107 children aged 0-1 in 2003-2004, and the kindergarten cohort (K-cohort) which follows 4983 children aged 4-5 in 2004.

[6] See Attanasio, Cattan, and Meghir (2022) for the review of research exploring the drivers of SES gaps in children's environments.

[7] For example, Boneva and Rauh (2018) document parental misperceptions about the timing of returns to early versus late childhood investment. Kiessling (2021) explore the relationship between perceived returns to parenting style and neighbourhoods and actual parental choices. Attanasio, Cunha, and Jervis (2019) show that mothers underestimate the returns to their investments in children and beliefs about returns predict actual investment choices.

[8] For example, using an experimental setting Dizon-Ross (2019), Doss et al. (2019), and Bergman (2021) show that mothers who update their beliefs about their children's academic progress also update choices of educational inputs. Bergman (2021) also shows that correcting mothers' misperceptions about children's learning efforts lead to improved students' learning efforts according to teachers' evaluations.

In this paper, I use the data for children between 4 and 9 years old, however, the most recent waves of the survey follow the kindergarten cohort till the age of 18-19 and the baby cohort till the age of 14-15.

The survey has four features that allow me to investigate the size and impact of teachers' reference group bias. First, it contains a rich set of objective child development measures obtained during the interview that can be matched to assessments by teachers and parents. Second, the sample is clustered at the neighbourhood level, which allows me to construct a measure of the local environment by matching children from the same neighbourhood. Third, it tracks the dynamics of multiple measures related to children's home and school environments. Finally, the survey collects a comprehensive set of information about family demographic and educational composition, family income, labour market outcomes, and neighbourhood characteristics.

2.1 Measures of child development

The survey collects information about child development from three sources: interviewers observing children during face-to-face surveys, teachers, and parents. First, trained interviewers use tests and direct observations to assess children's cognitive development and non-cognitive development. Second, children's teachers are invited to evaluate children's progress and their classroom environment. Finally, children's primary caregivers, mainly mothers, are asked to evaluate children's development and environment during face-to-face interviews.

I use direct observations of children's non-cognitive skills recorded by interviewers during the in-person household visit to construct a measure of the children's non-cognitive development. These observations are available when children are 4-5 and 8-9 years old. The face-to-face part of the interview lasts 1 - 2.5 hours with and without the parent present, giving interviewers a chance to observe children during a variety of interactions. The interviewers evaluate children's non-cognitive development across three dimensions: negative response, focus during the cognitive test, and positive response. The negative response includes fussing, pouting, whining, crying, and vocal/physical expressions of anger. Persistent loss of temper and aggressive behaviour are the symptoms of disruptive behaviour disorders in children.[9] The second dimension

[9] See Centers for Disease Control and Prevention https://www.cdc.gov/ncbddd/adhd/conditio

is the degree to which the child was able to sustain an interest in cognitive tasks. According to the Australian Psychological Society, difficulty concentrating and staying focused is the main symptom of ADHD in young children.[10] The third dimension is the degree of positive response by children, which includes smiling, laughing, or sounding excited, happy, or pleased.

These measures have three important advantages useful for quantifying the reference bias in teachers' assessments. First, all interviewers were trained by psychologists to conduct the evaluations consistently and went through practice interviews with parents and children. Second, interviewers relied on an objective scale to count the number of times and intensity of the children's responses. For example, the interviewers choose from 5 options while evaluating the children's negative or positive response: none displayed, 1-2 brief displays, 3 or more brief displays, 1-2 intense, heightened or prolonged displays, 3 or more intense, heightened or prolonged displays.[11] These two features ensure that the interview measures of non-cognitive development are objective and free from reference group bias. Finally, the interviewer's evaluations are available for a nationally representative sample of children. Observations of children's non-cognitive development did not add to the time or cost of the interview since the children were not asked to go through any additional testing. Therefore, the design of the interviewers' evaluation of the children's socio-motional development in LSAC allowed for large-scale direct evaluations.[12]

I summarize the children's non-cognitive development measured during the interview with the first principal component of the three age-standardized objective measures of non-cognitive skills. I age-standardize the interview non-cognitive score to make it comparable in scale to other skill measures used in the analysis. Figure 1 illustrates the distribution of the interviewers' observations when children were 4-5 years old. While children's degree of positive response varied a lot between children, 62 percent of children were both constantly focused during the interview and did not show any negative responses. Figure A.1 shows that the distribution of the interview non-cognitive score is skewed to the right, with many children receiving maximum or

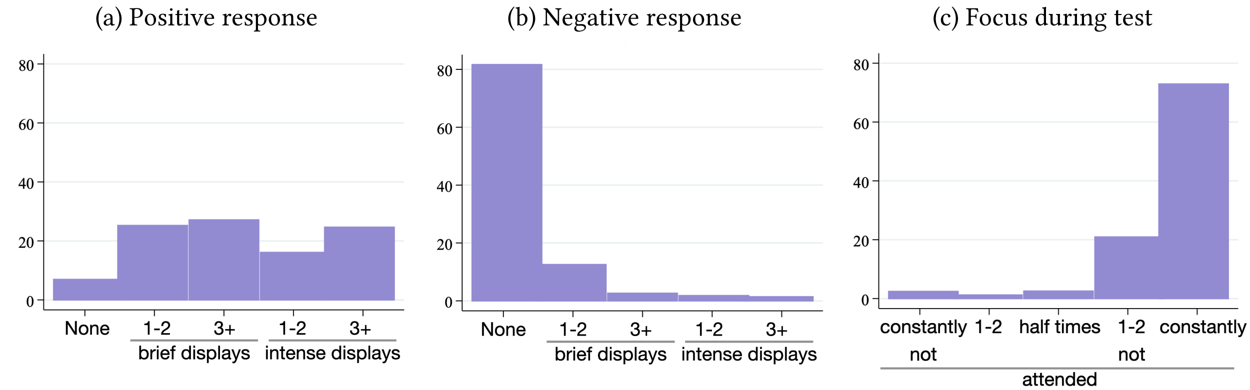
ns.html

[10]See https://psychology.org.au/for-the-public/psychology-topics/adhd-in-children.

[11]The 5 options for the degree of children's focus include: Constantly did not pay attention; Typically did not pay attention, attended in 1-2 instances; Did not pay attention half the time; Typically paid attention, but attention wandered in 1-2 instances; Constantly paid attention/concentrated.

[12]By contrast, evaluations of non-cognitive skills by psychologists who directly observe children are less common in large-scale survey datasets, as these types of evaluations are often more resource-intensive and time-consuming.

Figure 1: Non-cognitive development observed during the interview at ages 4-5



Notes: The figure displays histograms of the interviewer's records evaluating the non-cognitive skills of children ages 4-5 during the interview. Panel a) records the degree of positive response from the child, where the positive response includes smiling, laughing, or sounding excited, happy or pleased. Panel b) records the degree of negative response from the child, where the negative response includes fussing, pouting, whining, crying and vocal or physical expression of anger. Panel c) records the degree of the child's focus during the PPVT cognitive test.

near-maximum scores and a long left tail with children demonstrating non-cognitive problems with varying degrees of frequency and intensity.

The receptive language dimension of cognitive skills is measured during the interview by the short form of the Peabody Picture Vocabulary Test (PPVT). It is a standard age-adapted measure of children's receptive vocabulary and knowledge of spoken words. This test is commonly used in the literature as a measure of children's cognitive skills (Fiorini and Keane, 2014, Nicoletti and Tonei, 2020). Another measure of cognitive skills that is available in the survey is the Who Am I assessment (WAI) which is used for children aged 4–5 to measure the general cognitive abilities needed for beginning school. This measure tests receptive and expressive language and numeric abilities. I use it as an instrument to address measurement error in the PPVT measure.

A common critique of standardized tests of cognition is that they provide context-dependent measures of children's academic abilities. They depend on students' effort, motivation, test-taking abilities, and a range of other factors (Heckman and Kautz, 2012). The measure of children's non-cognitive skills obtained from a limited set of interactions between interviewers and children is likely to have similar limitations. By contrast, teachers' assessments are enriched by teachers interacting with children in multiple environments and learning about children's history, family, and community. While it is unlikely that the interview scores capture all the relevant information about children's development, they provide measures of non-cognitive development and cognition that are both free from reference bias and reflect variation in children's skills that predicts future outcomes.[13]

2.2 Measures of perceptions about child development

An important advantage of the LSAC survey is the availability of teachers' evaluations of children's developmental deficits. When children are 4-5 years old, teachers' beliefs about a child's skill relative to children of the same age are elicited through the teachers' questionnaire. Teachers are asked to evaluate the child's developmental level compared to other children of a similar age in several dimensions, including socio-emotional development (e.g. adaptability, cooperation, responsibility, self-control) and receptive language ability (e.g. understanding, interpreting, and listening).[14] These perceptions correspond to the same dimensions of child development that are measured by the interviewer, which allows me to match the measured dimension of development to teachers' evaluations of it and quantify the bias in teachers' assessments. The results of assessments performed during the interviews are unknown to teachers and parents. Teachers can rank the child as much more competent than others, as competent as other children, less competent than others, and much less competent than others. I construct a binary variable for the teacher indicating developmental delay in a particular dimension of the child's skills if the teacher responds that the child is less or much less competent than other children in that dimension. Only 15 percent of teachers indicate deficits in children's cognitive development and over 20 percent of teachers indicate deficits in children's non-cognitive development (see Appendix Table A.1).

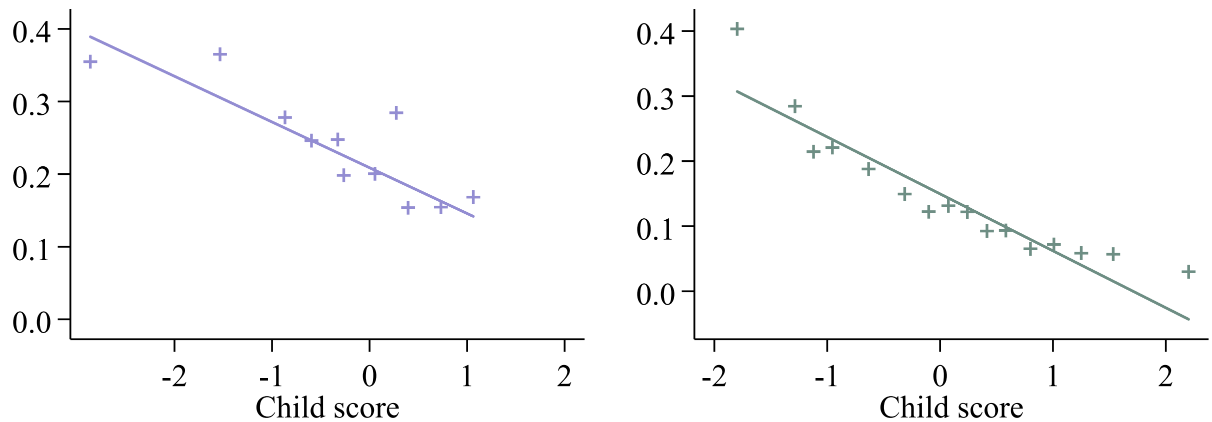
Figure 2 summarizes the relationship between teachers' perceptions and children's development scores measured during the interview. Panel (a) plots the average share of teachers reporting non-cognitive deficits against children's non-cognitive interview scores. The negative

[13] For example, children's cognitive and non-cognitive scores at ages 4-5 are associated with a lower likelihood of children repeating a grade by the age 12-13 and higher Grade 9 national test scores in reading and numeracy (see Appendix D).

[14] The exact question is "Rate how this child was compared with other children of a similar age, over the past few months."

Figure 2: Children's interview development scores and teachers' perceptions at ages 4-5.

(a) Share of teachers reporting non-cognitive delays (b) Share of teachers reporting cognitive delays



Notes: Panel (a) displays binned scatterplots of teachers' non-cognitive deficit recognition conditional on children's non-cognitive interview development score. Panel (b) displays binned scatterplots of teachers' cognitive deficit recognition conditional on children's cognitive interview development score.

slope implies that teachers' perceptions are informed by children's non-cognitive development measured during the interview. In particular, children who received higher non-cognitive development scores during the interview are less likely to be perceived as having non-cognitive delays by their teachers.[15] Similarly, Panel (b) shows that teachers are less likely to report cognitive delays for children who received higher cognitive interview scores. Even though the estimated relationship between the measured development and the probability that teachers indicate delays is negative, my analysis will show that this relationship is distorted by the reference bias in teachers' judgement.

The survey also collects mothers' perceptions about delays in their children's non-cognitive development. Mothers are asked to evaluate whether the child is more difficult compared to other children of similar age. Mothers can respond that the child is easier than average, about average, or more difficult than average. I assess whether mothers can identify non-cognitive developmental delays in their children if they perceive their child as more difficult compared to other children of similar age. Only 7 percent of mothers consider their child to have a delay in non-cognitive development (see Appendix Table A.1).

[15]Panel (a) also reflects that the distribution of interview non-cognitive scores is skewed to the right.

Importantly, measures of teachers' and mothers' perceptions evaluate their beliefs about children's development relative to the whole population of similar-aged children. Deficit identification by mothers and teachers, therefore, should involve a comparison of children's non-cognitive and cognitive skills to age-specific developmental milestones, as perceived by teachers and mothers, but not a comparison relative to children's class or grade level within their school or neighbourhood. When asking teachers or mothers to compare children's development against other children in the group, class, or grade level, the questionnaire explicitly provides the reference group.[16] Moreover, Kinsler and Pavan (2021) show that when evaluating cognitive skills mothers respond differently when asked to compare their child against other children of a similar age and other children in their child's class.

An important advantage of the LSAC for exploring the relationship between mothers' and teachers' perceptions is that when children are 8-9 years old, mothers are asked whether the school has contacted them about their children's behaviour within the last 12 months. This indicator allows me to go beyond estimating the potentially bidirectional association between mothers' and teachers' perceptions about children, and evaluate how mothers update their beliefs when they are informed by schools about their children's non-cognitive deficits.

2.3 Measures of school and home environments

I also explore the role of deficit recognition for school and family investments, as well as parental attitudes. The survey collects a wide variety of information about the home environment, such as parental and school investment into remedial services like behavioural, speech, and learning therapy, parenting styles, and tutoring. These measures allow me to explore the role of teachers' and mothers' perceptions in choices of inputs that have been shown to matter for child development by previous research (Cunha, Heckman, and Schennach, 2010, Caucutt, Lochner, Mullins, and Park, 2020, Del Boca, Flinn, and Wiswall, 2014, Fiorini and Keane, 2014).

The school- or neighbourhood-based investment measures include two types of therapy, one directed at the treatment of cognitive skills (learning or speech therapy), and another directed at the treatment of non-cognitive skills (behavioural or psychological therapy). LSAC asks

[16] For example, in the same questionnaire teachers' are asked: "During organized physical activities for your group, how does this child compare with other children in the group in terms of the level of physical activity?"

teachers and mothers whether children have used additional school or community services that can allow children with delays to catch up. I create a binary variable indicating that the child has used behavioural or psychological therapy if teachers report that the child used behaviour management programs or had a psychological assessment while being in their care or if mothers report that, in the last 12 months, they have used a guidance counsellor or other psychiatric or behavioural services for the child. I create a binary learning or speech therapy variable equal to one if teachers indicate that the child has used speech therapy or learning support while in their care and if mothers indicate that the child has used speech therapy in the last 12 months. At ages 4-5, around 4 and 14 percent of children are reported to have received non-cognitive and language therapy, respectively (see Appendix Table A.2).

Measures of family investments and parental attitudes include LSAC-constructed scores measuring warmth and anger of mothers' parenting style, weekly tutoring sessions, expectations about children's future educational achievements, and other measures of quality time with household members including the total time, and weekly times household members spent reading, playing, and drawing with the child, etc.[17]

2.4 Measures of the local environment

To measure the local environments of children, I exploit information on the current household location. I use postcodes as the geographic unit defining neighbourhoods. Crucially, the sampling design of the LSAC survey allows the grouping of children together based on their current postcode.[18] There are a total of 2,644 postcodes in Australia. The sample in the first wave of the survey includes children from 409 postcodes representing all Australian territories with an average of 37 children per postcode. For example, in Sydney and Perth children from 93 and 35 postcodes, respectively, were selected to participate in the survey. The availability of multiple child observations per neighbourhood and the sampling design, which is representative of the population of Australian children, allows me to compute neighbourhood child development levels.

[17] See Appendix A.4 for the summary of family investment measures.

[18] A sample was selected to be representative of all Australian children in the selected age cohorts. It was drawn using the two-stage stratified sampling procedure, with the first stage including a selection of postcodes to ensure proportional geographic representation for Australian territories, and the second stage selection of children from these postcodes.

To analyze the effect of local environment on the identification of children's developmental delays, I calculate it using the average levels of non-cognitive and cognitive development of other children living in a child's neighbourhood. To quantify the reference bias in teachers' assessments, I need objective measures of average skills across neighbourhoods. For each child, I compare children of a similar age in both cohorts living in the same postcode. I construct the average neighbourhood score in two steps. First, I de-mean interview development measures by year and age group. Second, I use de-meaned scores excluding children's own scores to construct the leave-one-out average neighbourhood score if measures for at least 10 children other than the study child were available in a given neighbourhood across both cohorts. The measure, therefore, allows the distribution of development levels to shift in a parallel fashion across time.[19] I age-standardize the neighbourhood's average levels of development to make the scale comparable to individual assessments.

3 Conceptual framework: The role of bias in teachers' perceptions

This section describes a conceptual framework in which reference group bias in teachers' perceptions can lead to a cascading effect on the children's environment inspired by the setting in Kinsler and Pavan (2021). Consider a child of age t with a development level D_t. During the interview, this development level is evaluated by a psychologist-trained interviewer who assigns a continuous measure D^I_{it} so that

$$D^I_{it} = D_{it} + \mu^I_{it}, \text{ s.t. } \mu^I_{it} = \Theta^I_{it} + \epsilon^I_{it}, \tag{1}$$

where μ^I_{it} summarizes potentially unobserved factors that can affect the objective interview measures, ϵ^I_{it} is a mean-zero iid measurement error, and $\Theta^I_{i,t}$ is an idiosyncratic interview day shock which can be correlated across different interview development measures. For example, this shock can capture the effect of the child being distracted or helped by parents during the inter-

[19] See Appendix F for the discussion of modifications to the construction of the average neighbourhood development.

view and hence applying little effort during evaluations.

Teachers aim to identify developmental delays in children, $T_{it} = \{0,1\}$. To do so, they compare children's development against perceived age-specific developmental benchmarks. If these perceived standards of development are affected by the level of development of other children in the neighbourhood, teachers' deficit recognition depends both on a child's own development level and on the average development level in the neighbourhood, $\bar{D}^N_{it}$. Thus, teachers' perceptions about children's deficits relative to other children of the same age are

$$T_{it} = F^T(D_{it}, \bar{D}^N_{it}, X^T_{it}) + \mu^T_{it}, \ \text{ s.t. } \ \mu^T_{it} = \Theta^T_{it} + \epsilon^T_{i,t}, \tag{2}$$

where X^T_{it} are variables related to children's development and perceptions that are observed by both interviewers and teachers. These variables can affect perceptions conditional on actual children's development levels, for example, teachers can perceive children from lower socio-economic status (SES) families to be more prone to non-cognitive delays, and neighbourhoods with low child development might have a higher number of these families. Here, μ^T_{it} summarizes factors unobserved by the interviewer that can affect teachers' perceptions, ϵ^T_i represents an iid error term, and $\Theta^T_{i,t}$ represents elements not captured by interview measures but potentially related to children's or neighbourhood's average development levels and perceptions. If teachers' benchmarks for what constitutes healthy development depend on the local environment $\bar{D}^N_{it}$ in a systematic way, then teachers' evaluations of children's developmental deficits are distorted compared to the objective developmental milestones for the population of children of the same age.

Mothers' perceptions about children's developmental deficits M_{it} depend on children's true development levels and teachers' identification of children's developmental delays

$$M_{it} = F^M(D_{it}, T_{i,t}, X^M_{it}) + \mu^M_{it}, \quad \text{ s.t. } \ \mu^M_{it} = \Theta^M_{it} + \epsilon^M_{it}, \tag{3}$$

where X^M_{it} are observed variables potentially correlated with children's development and perceptions. The shifter $\Theta^M_{i,t}$ can include elements of idiosyncratic perceptions of mothers like over-optimism or the lack of involvement, which can be correlated with children's development or teachers' perceptions. For example, teachers might communicate differently with mothers who are uninvolved, and these mothers can also be more likely to have children with lower skill levels. If teachers' assessments are biased, and mothers' perceptions about children are affected by teachers' perceptions, then the bias in teachers' evaluations will be transmitted to mothers.

Teachers' and mothers' perceptions play a critical role in children's environment (Dizon-Ross, 2019). Formally, the school-based investments, $I^S_{i,t}$, are determined by teachers' and mothers' perceptions, while family-based investments, $I^F_{i,t}$, are determined by mothers' perceptions

$$I^S_{i,t} = F^S(M_{it}, T_{it}, X^S_{it}) + \mu^S_{i,t} \text{ s.t. } \mu^S_{i,t} = \Theta^S_{i,t} + \epsilon^S_{i,t},$$

and

$$I^F_{i,t} = F^F(M_{it}, X^F_{it}) + \mu^F_{i,t} \text{ s.t. } \mu^F_{i,t} = \Theta^F_{i,t} + \epsilon^F_{i,t}.$$

Here, $\Theta^S_{i,t}$ can include idiosyncratic determinants of school investments like available resources and $\Theta^F_{i,t}$ can include unobserved determinants of family investment like habits. These shifters can be correlated with perceptions and investment choices, for example, uninvolved mothers can be less likely to recognize deficits and habitually invest less in their children. If family and school investments depend on how teachers and mothers perceive children's developmental delays, then distortions in perceptions about children have a cascading effect on children's environment, potentially leading to suboptimal investment strategies for children. The following sections describe my estimation strategy and elaborate on the results.

4 Teachers' perceptions and local environment

Teachers' and mothers' evaluations of children's non-cognitive skills are commonly used to compare the levels of development across different groups of children. For example, in Australia, the Australian Early Development Census surveys teachers of children ages 4-5 across the country to identify communities and institutions that are struggling to promote non-cognitive development in children. Similarly, research on child development commonly relies on teachers' or mothers' evaluations to compare levels of non-cognitive skills across children or across time (Attanasio, De Paula, and Toppeta, 2020, Chaparro, Sojourner, and Wiswall, 2020, Fletcher and Wolfe, 2016, Nghiem, Nguyen, Khanam, and Connelly, 2015). Any reference group bias in teachers' assessments is likely to distort these estimated differences in skills. This section quantifies the reference group bias in teachers' evaluations of child development using objective measures of child development and average neighbourhood child development levels.

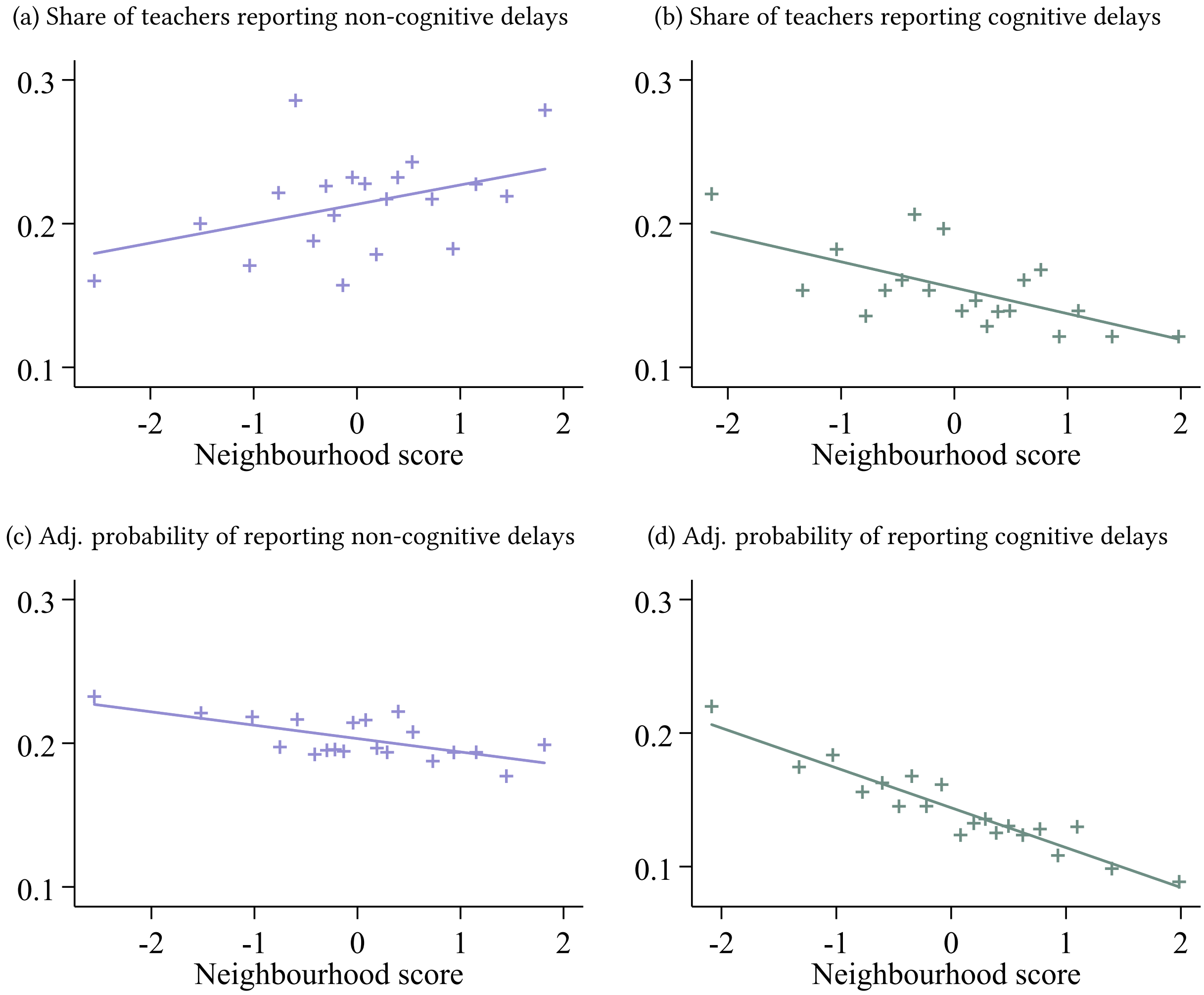
Children who live in relatively more advantaged areas are likely to have higher objective development levels. As teachers' perceptions about children's developmental delays are informed by children's development as was shown in Figure 2, it can also be expected that teachers in neighbourhoods with lower average levels of child development are more likely to report developmental delays in children. Panels (a) and (b) of Figure 3 plot the average shares of teachers reporting developmental delays against the average neighbourhood interview development scores. While Panel (b) shows that teachers report fewer delays in more cognitively developed neighbourhoods, Kinsler and Pavan (2021) and Elder and Zhou (2021) have shown that this negative relationship would have been stronger in the absence of the reference bias in teachers' perceptions. Importantly, the relationship between deficit recognition by teachers and neighbourhood average development levels is positive for non-cognitive skills. Teachers in neighbourhoods with higher average levels of non-cognitive development measured during the interview are actually more likely to report non-cognitive delays in children. This relationship can be driven by teachers having higher expectations about children's developmental milestones in neighbourhoods where the majority of children are developmentally on track.

To see whether teachers in neighbourhoods with lower levels of child development are less likely to report developmental deficits in children, I estimate a linear probability regression based on Equation (2). The dependent variable is a binary variable reflecting the teachers' recognition of the child's i developmental deficit, T_{it}, at age t. The child has a measured level of development, D^I_{it}, and the average neighbourhood level of development, $\bar{D}^N_{it}$, then

$$T_{i,t} = \beta^{T,N}\bar{D}^N_{i,t} + \beta^{T,D}D^I_{i,t} + \gamma_t^{T,X}X^T_{i,t}, \tag{4}$$

where X^T_i is a vector of control variables included in all specifications. It contains the child's gender, cohort, age in months, and an index for the socioeconomic status (SES) of the house-

Figure 3: Average neighbourhood development scores and teachers' perceptions at ages 4-5.



Notes: Panels (a) and (b) report the binned scatterplot of the shares of teachers reporting developmental delays conditional on the average neighbourhood development score. Panel (a) displays binned scatterplots of teachers' non-cognitive deficit recognition conditional on the average neighbourhood non-cognitive interview development score. Panel (b) displays binned scatterplots of teachers' cognitive deficit recognition conditional on the average neighbourhood cognitive interview development score.
Panels (c) and (d) report the binned scatterplot of predicted probabilities to report developmental delays adjusted for the reference bias according to the estimates reported in columns (2) and (4) in Table 1. The adjusted probabilities are constructed by predicting the probability while setting average postcode levels of cognitive and non-cognitive development to the mean (zero) value. Panel (c) displays binned scatterplots of the adjusted probability to report non-cognitive deficit conditional on the average neighbourhood non-cognitive interview development score. Panel (b) displays binned scatterplots of the adjusted probability to report cognitive deficits conditional on the average neighbourhood cognitive interview development score.

Table 1: Teachers' beliefs and neighbourhood child development levels.

	Non-cognitive delay		Cognitive delay	
	(1)	(2)	(3)	(4)
Neighbourhood non-cognitive score	0.024*	0.020*		0.014*
	(0.006)	(0.006)		(0.005)
Non-cognitive score	-0.057*	-0.043*		-0.030*
	(0.006)	(0.007)		(0.006)
Neighbourhood cognitive score		0.010	0.019*	0.016*
		(0.006)	(0.005)	(0.005)
Cognitive score		-0.052*	-0.086*	-0.082*
		(0.007)	(0.006)	(0.006)
N	5520	5258	5270	5254

Notes: Linear probability regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. Standard errors are clustered at the postcode level. Significance level: * 5%.

hold derived in LSAC, calculated as a weighted composite of parental income, education, and occupational prestige (Baker, Sipthorp, and Edwards, 2017).[20]

Table 1 shows the estimates of the linear probability regression specified by Equation (4). Columns (1) and (3) show the results when only one dimension of development is analyzed. Dependent variables include the interview measure and the neighbourhood average of the development level. The estimates show that teachers are less likely to indicate delays in the development of children who are further ahead according to the interview assessments, even though they do not observe the interview evaluations of non-cognitive or cognitive development. Columns (2) and (4) show the results taking into account both dimensions of children's development. Teachers' evaluations of child development in both developmental dimensions are affected by both the children's non-cognitive and cognitive development levels. So, children who have higher non-cognitive skills are less likely to be identified as having delays in cognitive, and children who have better cognitive skills are less likely to be identified as having non-cognitive delays.

Most importantly, Table 1 provides evidence that the likelihood that teachers identify de-

[20]The results are similar if I control for mothers' age, number of siblings, family income, mothers' education, mothers' marital status, and mothers' employment status instead of the indicator for SES. See Appendix F for results with the full set of SES controls. Here I report the more parsimonious specification.

velopmental delays in children increases with the average level of non-cognitive development in their neighbourhood. This holds for measures of non-cognitive and cognitive deficits. In neighbourhoods where children, on average, have lower interview non-cognitive scores, teachers are less likely to report non-cognitive and cognitive developmental delays conditional on objective measures of children's skills. Further, the average level of cognitive skills in the neighbourhood affects teachers' recognition of children's cognitive delays, consistent with the findings of Kinsler and Pavan (2021) and Elder and Zhou (2021). However, it does not affect the recognition of delays in non-cognitive development by teachers.

Columns (2) and (4) of Table 1 also illustrate the limited effect of omitted variable bias on the estimates reported in Columns (1) and (3). This bias can be driven by the Θ_{it}^{T} in Equation (2), which is potentially unobserved by the interviewer but related both to the children's development and the teachers' assessments. For example, accounting for the children's cognitive skills in Column (2) in addition to their non-cognitive development reduces the estimates of the role of children's non-cognitive scores compared to those reported in Column (1) but not in a statistically significant way.

The estimates reported in Table 1 are based on several assumptions. First, I assume linearity of the functional form $F^{T}(.)$ in Equation (2). Second, I assume that there are no idiosyncratic interview days shocks or measurement error (Θ_{it}^{I} or ϵ_{it}^{I} in Equation (1)) since the children's true development is measured during the interview. Table 2 shows that relaxing these assumptions does not change the conclusions from Table 1 that teachers' perceptions depend systematically on the average level of development in the neighbourhood. For example, the estimates reported in Columns (2) and (4) of Table 1 are robust to relaxing the assumption of the linearity of $F^{T}(.)$. Columns (1) and (4) in Table 2 report the average marginal effects of the logistic probability model. The effect of the average neighbourhood non-cognitive development on deficit recognition by teachers remains positive and statistically significant.

The estimates of Equation (4) reported in Table 1 can also be affected by the idiosyncratic shocks to the interview development scores (Θ_{it}^{I}) described in Equation (1). To control for the potential effect of idiosyncratic shocks in the interview development scores (Θ_{it}^{I}) I estimate Equation (4) with added controls that proxy for variation in children's effort during the interview. These controls are available for the Baby cohort in LSAC, which characterize the behaviour of parents and siblings during the cognitive test. I account for indicators of whether the parent and sibling were not present in the room, present at a distance, observed the child, encouraged the child, or interfered with the tests. I also include a measure of children's sleeping problems and a set of indicators for the month of the interview. Finally, I control for a range of characteristics of children's teachers and classrooms to proxy for the potential selection of higher-quality teachers and programs with better conditions for deficit recognition into more developed neighbourhoods.[21] I control for whether the teacher has a university degree (versus diploma or certificate), whether the child is attending daycare (versus kindergarten or preschool), the age range of the children's class reported by the teacher, the ratio of children to qualified staff, and indicators measuring teachers' experience in childcare (0-5 years and 6-10 years versus more than 10 years).[22] Columns (2) and (5) in Table 2 show that these controls have little effect on the estimates of both the role of the individual non-cognitive score and the average neighbourhood non-cognitive score.

The estimates of Equation (4) reported in Table 1 also abstract from the potential measurement error in the interview development scores (ϵ_{it}^{I}) described in Equation (1). To explore the role of the measurement error, I estimate a TSLS version of the regression in Equation (4) using one of the three non-cognitive interview measures as the main measure of non-cognitive development and instrumenting it with the other two measures. I use children's focus during the cognitive test as the main measure of non-cognitive skills and instrument it with the degrees of positive and negative response during the interview. Similarly, I use the PPVT score as the main measure of cognitive skills and instrument it with the WAI score, with all measures age-standardized. Columns (3) and (6) in Table 2 illustrate the results adjusted for the potential measurement error in children's interview assessments. The measurement error has a substantial effect on the estimates, with the bias distorting estimates of both the children's individual development and the role of the neighbourhood's average development levels towards zero. Adjusting for measurement error strengthens the estimated effects of both children's own development and average neighbourhood development on the teachers' assessments.

[21]Subsection 4.1 will explore the role of teachers' qualifications and classroom characteristics for developmental deficit recognition in greater detail.

[22]See the summary of additional controls in Appendix A.5.

Table 2: Robustness checks

	Non-cognitive delay			Cognitive delay		
	(1)	(2)	(3)	(4)	(5)	(6)
	Logit Avg. Marg. Effect	Extra control	Meas. error adj.	Logit Avg. Marg. Effect	Extra control	Meas. error adj.
Neighbourhood	0.019*	0.023*	0.050*	0.012*	0.021*	0.025*
non-cognitive score	(0.006)	(0.009)	(0.013)	(0.005)	(0.007)	(0.010)
Non-cognitive score	-0.037*	-0.033*	-0.386*	-0.023*	-0.030*	-0.161
	(0.006)	(0.011)	(0.112)	(0.005)	(0.010)	(0.087)
Neighbourhood cognitive	0.010	0.007	0.003	0.017*	0.011	0.030*
score	(0.006)	(0.010)	(0.018)	(0.005)	(0.009)	(0.013)
Cognitive score	-0.050*	-0.074*	-0.045	-0.077*	-0.089*	-0.155*
	(0.006)	(0.011)	(0.067)	(0.005)	(0.011)	(0.052)
N	5258	1939	5215	5254	1939	5211

Notes: All columns control for children's gender, cohort, and age in months, household socioeconomic status (SES) index. Columns (1) and (4) report marginal effects of logistic probability model estimates. Columns (2) and (5) report the linear probability regression with added controls for sleeping problem intensity, age of the youngest in the class, age of the oldest in the class, whether the child attends daycare, the behaviour of parents and siblings during the test, the month of the interview., and teachers' and class characteristics. Standard errors are clustered at the postcode level. Columns (3) and (6) report linear probability model TSLS estimates adjusted for the measurement error in children's development. Instruments for focus during cogn. test: positive and negative behaviour during the interview. Instrument for PPVT: WAI. Standard errors are clustered at the postcode level. Significance level: * 5%

Next, I predict the gaps in teachers' deficit recognition between less- and more-advantaged areas which would manifest if teachers compared children against the same developmental standard. To do so, I predict probabilities of reporting delays based on estimates reported in columns (2) and (4) of Table 1 while setting the average neighbourhood cognitive development levels equal to the population mean (zero) values. Panels (c) and (d) plot these adjusted probabilities against the average neighbourhood development levels. Panel (c) shows that adjusting for the reference bias in teachers' perceptions about children's non-cognitive delays according to estimates reported in column (2) of Table 1 changes the sign of the relationship between teachers' perceptions and average neighbourhood non-cognitive development level and makes it negative. Teachers in neighbourhoods at the top quartile of average children's non-cognitive development are 1.5 percentage points more likely to report delays in children's non-cognitive development compared to teachers in bottom-quartile neighbourhoods. By contrast, the adjusted probability of reporting non-cognitive delays is 2.5 percentage points lower for teachers in top-quartile neighbourhoods compared to teachers in bottom-quartile neighbourhoods. Moreover, adjusting for the reference bias in teachers' assessments while accounting for the measurement error according to the estimates reported in column (3) of Table 2 increases the magnitude of this gap to 10 percentage points. Similarly, teachers in the top quartile of neighbourhoods in terms of average cognitive development are only 3.7 percentage points less likely to report cognitive delays. However, adjusting for the reference bias using the estimates reported in column (4) of Table 1 increases the magnitude of this gap to 7 percentage points. Using the estimates adjusted for the measurement error further increases the magnitude to 10.6 percentage points.

4.1 Teachers' quality and deficit recognition

It is important to understand what factors can help reduce the bias in teachers' evaluations. An extensive literature concerned with the role of teacher quality for student outcomes explores the role that teacher qualifications as well as classroom characteristics play in determining student progress.[23] I explore whether teacher qualifications, as measured by education or experience, or classroom characteristics, as measured by the type of childcare arrangement, class size or age composition, help teachers identify children with low levels of development.

[23] See Manning, Wong, Fleming, and Garvis (2019) for review of research on the role of teachers' qualifications.

The majority of children ages 4-5 in the sample attend preschool or kindergartens, while 26 percent of them attend daycare (see Appendix A.7). In Australia, early childhood educators without a university degree can be recognized as "equivalent early childhood teachers" if they go through the process of certification. Overall, 37 percent of teachers in the sample do not have university training.[24] This allows me to evaluate whether teachers with different levels of training differ in their evaluations of children's delays. Figure 4 shows the probability of teachers documenting cognitive and non-cognitive delays in children conditional on their qualifications. Panel (a) of Figure 4 shows that teachers with a university degree are 4 percentage points more likely to indicate delays in non-cognitive development and 3.6 percentage points more likely to report delays in cognitive development compared to teachers with certificates and diplomas. Moreover, Panel (b) of Figure 4 shows that the likelihood of teachers' deficit recognition is relatively stable across teachers' careers. The likelihood of indicating delays in non-cognitive development has an inverted U-shape, peaking when teachers have 11 to 15 years of experience. By contrast, for cognitive development inexperienced teachers are most likely to report delays. These differences in the rates of reporting can be driven by some teachers being more effective in recognizing low development in children, or by teachers reporting more deficits overall.

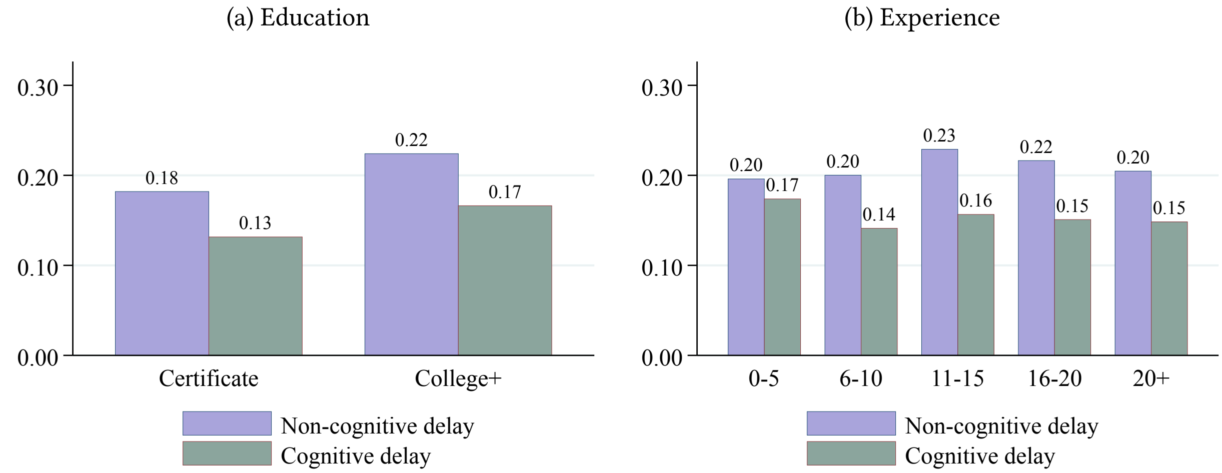
Table 3 explores whether teacher or classroom characteristics matter for teachers' recognition of low development levels. For each dimension of development I split the sample into subsamples with low- and high- measured levels of development based on the objective interview development scores. I assume that children have low measured non-cognitive (cognitive) development if their interview non-cognitive (cognitive) score is below the sample median for children ages 4-5. Intuitively, children in the subsample with the measured development below the median are more likely to have developmental deficits. Next, I estimate the linear probability regression separately for two subsamples, $J = \{H, L\}$:

$$T_{i,t} = \beta^{J,V} V_{i,t}^T + \gamma_t^{J,X} X_{i,t}^T,$$

where X_i^T is a vector of control variables included in all specifications. It contains the child's gender, cohort, age in months and an index for the socioeconomic status (SES). Here $V_{i,t}^T$ includes

[24] See Appendix A.6 for the summary of teachers' qualifications.

Figure 4: Share of teachers indicating delays for children ages 4-5 by teachers' characteristics.



Notes: Panel (a) shows the average share of teachers reporting delays in cognitive and non-cognitive development by teacher education. Panel (b) shows the average share of teachers reporting delays in cognitive and non-cognitive development by teacher work experience in childcare.

observed teacher and classroom characteristics that can be associated with the quality of early childhood education.[25] In particular, I account for whether the teacher has a university degree (versus diploma or certificate), whether the child is attending daycare (versus kindergarten or preschool), the age range of children in class reported by the teacher, the ratio of children to qualified staff, and indicators measuring teachers' experience in childcare (0-5 years and 6-10 years versus more than 10 years).

Teacher education is positively associated with deficit recognition for subsamples of children with low measured levels of development. Teachers with a bachelor's or postgraduate university degree are 6.2 percentage points more likely to report deficits in children's non-cognitive skills for the subsample of children with low measured non-cognitive development compared to teachers with diplomas or certificates. They are also 5.1 percentage points more likely to report cognitive delays for the subsample of children with the low measured cognitive development. As can be expected, this relationship is not evident for children with high measured development, where more extensive training should not increase the likelihood of deficit recognition.

[25]These characteristics include teacher and classroom qualities used as controls for the estimation reported in columns (2) and (5) of Table 2.

Table 3: Teachers' and classroom characteristics and deficit recognition

	Non-cognitive delay		Cognitive delay	
	(1)	(2)	(3)	(4)
	Non-cogn. score low	Non-cogn. score high	Cogn. score low	Cogn. score high
Teacher college+	0.062*	0.023	0.051*	0.009
	(0.018)	(0.015)	(0.015)	(0.011)
Child attends daycare	-0.036	-0.018	-0.024	0.002
	(0.022)	(0.018)	(0.018)	(0.014)
Teaching experience 0-5 years	-0.042	0.012	0.007	0.032
	(0.022)	(0.020)	(0.020)	(0.018)
Teaching experience 6-10 years	0.009	-0.015	-0.022	0.011
	(0.021)	(0.017)	(0.018)	(0.014)
Age of youngest in class	0.001	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Age of oldest in class	-0.001	0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.000)
Children to qualified staff ratio	-0.000	0.002	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
N	2899	2847	2771	2749

Notes: Linear probability regression for high-development (measured development higher than the median) and low-development (measured development lower than the median) samples. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. Standard errors are clustered at the postcode level. Significance level: * 5%.

Teachers' experience or other classroom characteristics do not have a robust association with deficit recognition. For the subsample with low measured non-cognitive development, inexperienced teachers with less than 5 years of experience are 4.2 percentage points less likely to report delays in children's non-cognitive development compared to teachers with more than 10 years of experience, however, the estimated effect is not statistically significant. A further increase in experience is not associated with gains in the probability of deficit recognition. This may reflect the fact that teachers persistently sort into low-development or high-development areas. In that case, working extra years does not expand the teacher's reference group to include children with different levels of skills. By contrast, being exposed to university training can improve the ability of teachers to correctly identify children with delays even in neighbourhoods with low average child development levels. Similarly, the class size as measured by the children-to-qualified-staff ratio or the age composition of the class do not have a noticeable effect on deficit recognition.

Next, I explore how the reference bias differs between teachers with different education levels. Table 4 reports the estimates of Equation (4) estimated separately for the subsamples of teachers with certificates or diplomas, and teachers with undergraduate or postgraduate degrees. The regressions additionally control for the set of teachers' and classroom characteristics. While the estimated magnitudes of the reference bias do not vary substantially between teachers with different education levels, teachers with a university degree have a stronger association between children's own measured development and deficit recognition. This difference is particularly pronounced in the estimated coefficients for the interview measures of cognitive skills. One possible interpretation of this can be that the teachers who receive university training become more equipped to evaluate children's progress in receptive language. This can explain why teachers with university degrees are more likely to indicate delays for subsamples of children with low development levels.

Teachers' professional judgement plays a critical role in facilitating children's learning. For example, the Early Years Learning Framework published by the Australian Government Department of Education emphasizes the fundamental role that teachers' assessments play in the effective planning of children's learning, communicating about children's progress, determining

Table 4: Teachers' education and reference group bias

	Non-cognitive delay		Cognitive delay	
	(1)	(2)	(3)	(4)
	Certificate	College+	Certificate	College+
Non-cognitive score	-0.031*	-0.055*	-0.030*	-0.036*
	(0.011)	(0.009)	(0.011)	(0.008)
Neighbourhood non-cognitive score	0.016	0.025*	0.021*	0.016*
	(0.010)	(0.008)	(0.008)	(0.007)
Cognitive score	-0.030*	-0.067*	-0.060*	-0.098*
	(0.012)	(0.009)	(0.009)	(0.008)
Neighbourhood cognitive score	0.004	0.008	0.012	0.018*
	(0.010)	(0.009)	(0.008)	(0.008)
N	1725	2912	1722	2912

Notes: Linear probability regression for subsamples of teachers with certificates of diplomas or undergraduate and postgraduate degrees. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index, whether the child is attending daycare (versus kindergarten or preschool), the age range of the children's class reported by the teacher, the ratio of children to qualified staff, and indicators measuring teachers' experience in childcare (0-5 years and 6-10 years versus more than 10 years). Standard errors are clustered at the postcode level. Significance level: * 5%.

the extent to which all children are progressing toward their learning outcomes, and identifying children who need additional support (Australian Government Department of Education, 2009). The bias in teachers' evaluations is likely to distort all of these processes, generating a cascading effect on parental perceptions and children's learning environments. In the next sections, I shed light on this cascading effect of teachers' misperceptions by showing, first, that teachers' perceptions of child development affect parental perceptions, and, second, that the judgement of parents and teachers about child development alters children's learning environments.

5 The influence of teachers' perceptions on mothers

Teachers are often a key source of information for parents trying to monitor their children's progress. This section investigates whether biased teachers' perceptions about child development are likely to be transmitted to parents. To explore the connection between teachers' and mothers' perceptions, I utilize the measures of mothers' and teachers' perceptions of child non-cognitive development, as well as objective measures of child development available for ages 4-5 and 8-9.

I estimate a linear probability regression where the dependent variable is equal to one if the mother has indicated her child to be more difficult than other children of similar age (M_{it}).

$$M_{it} = \beta^{MD} D^I_{it} + \beta^{MT} T_{it} + \gamma^{MX} X^M_{it}, \tag{5}$$

where X^M_i - is a vector of controls including children's gender, cohort, age in months, the index for the socioeconomic status (SES) of the household, and mothers' depression levels. Some specifications additionally control for the lag of mother perceptions and mothers' level of involvement at school as reported by teachers when this measure is available. T_{it} is a measure of teachers' identification of children's developmental delays, where I explore different measures of teachers' perceptions for children ages 4-5 and for children ages 8-9.

Columns (1) and (2) in Table 5 show the relationship between mothers' deficit recognition and teachers' perceptions when children are aged 4-5. All variables for these specifications are measured when children are ages 4-5. I use teachers' non-cognitive or cognitive deficit recogni-

Table 5: Mothers' and teachers' perceptions

	Non-cognitive delay perceived by mother			
	(1)	(2)	(3)	(4)
	Ages 4-5	Ages 4-5	Ages 8-9	Ages 8-9
Teach.: Non-cognitive delay	0.096*	0.080*		
	(0.013)	(0.020)		
Teach.: Cognitive delay	0.042*	0.020		
	(0.015)	(0.021)		
School contacted about behavior			0.180*	0.115*
			(0.015)	(0.015)
Mother depression	0.028*	0.021*	0.023*	0.016*
	(0.004)	(0.006)	(0.003)	(0.003)
Non-cognitive score	-0.021*	-0.020*	-0.024*	-0.015*
	(0.005)	(0.007)	(0.004)	(0.004)
Cognitive score	-0.004	0.002	-0.012*	-0.005
	(0.004)	(0.006)	(0.004)	(0.003)
Lag mother: Non-cognitive delay		0.313*		0.510*
		(0.044)		(0.026)
Index: mothers' interactions with school				0.003
				(0.003)
N	4733	2228	7261	5561

Notes: Linear probability regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index, mothers' depression score. Standard errors are clustered at the postcode level. Significance level: * 5%

tion analyzed in Section 4 to measure teachers' perceptions. Mothers of children whose teachers indicate developmental delays in the non-cognitive dimension are indeed more likely to think that their children have non-cognitive delays. Column (2) additionally controls for the lag of maternal perceptions measured when children are ages 2-3.[26] The lag of maternal perceptions accounts for persistent elements of unobserved heterogeneity in perceptions that can be correlated with factors affecting the transmission of teachers' perceptions represented by Θ_{it}^{M} in Equation (3). This can be driven by different valuations of skills, over-optimism, lack of involvement, etc. If teachers indicate non-cognitive delays in children's development, mothers are 8 percentage points more likely to also indicate that their children are delayed in development. However, this association does not necessarily imply that teachers' perceptions transfer to parental perceptions about children's non-cognitive development since mothers' and teachers' perceptions about children's development are likely to affect each other. For example, mothers of children diagnosed with ASD or ADHD can contact teachers to inform them about the developmental needs of their children.[27]

To address the reverse causality problem, I take advantage of the LSAC question asking mothers whether, in the last 12 months, the school has contacted them because their child has behaved poorly at school when children are 8-9 years old. This represents a direct information transfer from teachers to parents, which signals delays in children's non-cognitive development. For children ages 8-9 columns (3) and (4) of Table 5 show that mothers update their beliefs about children's non-cognitive delays after being contacted by the school. Column (4) additionally controls for the lag of mothers' perceptions measured when children were ages 6-7 and the level of mothers' involvement at school as reported by teachers.[28] These variables proxy for differences in incentives of schools to contact mothers with different levels of engagement. For example, schools might not need to contact mothers who maintain constant communication with teachers. Similarly, teachers might be more reluctant to contact mothers who seem uninvolved in

[26] Since the first age of observations for the kindergarten cohort in LSAC is 4-5, this regression can only be estimated for the baby cohort followed from ages 0-1.

[27] From columns (1) and (2) it also follows that mothers with more depressive symptoms are more likely to perceive their child to be more difficult than other children, similar to the findings of Del Bono, Kinsler, and Pavan (2020).

[28] The index of mothers' involvement at school is computed by LSAC based on subquestions measuring parents' interactions with school. Parents' interactions include contacting the teacher, visiting the child's class, volunteering in class, helping in school, attending parent-school committees, raising funds, and participating in other activities.

children's education. Column (4) reports that being contacted by the school in the last 12 months increases the probability that mothers perceive their children to be more difficult than others by 11.5 percentage points.

This section shows that teachers' judgement affects parental perceptions about children's non-cognitive skills. While teachers' evaluations of academic progress are expected to inform parental knowledge (Dizon-Ross, 2019, Doss, Fahle, Loeb, and York, 2019), the role of teachers in parental learning about children's non-cognitive development is less understood. The estimated impact of communication between teachers and parents in the presence of bias in teachers' assessments implies that parental perceptions are affected. Any distortions can be transmitted to child environments through the optimal investment choices of teachers and parents. The next section studies the relationship between non-cognitive deficit recognition in children and investment into skills, educational aspirations and parental attitudes.

6 Teachers' perceptions and school environment

Both families and schools can undertake investments in children who are falling behind. This section investigates the extent to which the bias in teachers' recognition of child developmental delays can affect children's environment through family and school investments. I focus on the neighbourhood- or school-level investment into different types of child therapy, as well as a range of family investments that have been shown to have an important role in child development, including parenting style and attitudes, time investment, and extra cost activities like tutoring (Cunha, Heckman, and Schennach, 2010, Caucutt, Lochner, Mullins, and Park, 2020, Del Boca, Flinn, and Wiswall, 2014, Fiorini and Keane, 2014).

One compensatory mechanism available to schools and families is therapy. For children aged 4-5, I explore the role of deficit recognition in their participation in two types of therapy: one focused on cognitive skills (learning or speech therapy) and another focused on non-cognitive skills (behavioural or psychological therapy). To understand the role of teachers' perceptions in the uptake of therapy, I estimate the linear probability regression where the dependent variable is equal to one if the child is getting therapy ($S_{it} = \{0,1\}$), and the independent variables include the teachers' and maternal identification of the children's developmental delays (T_{it} and M_{it})

$$S_{it} = \beta^{ST} T_{it} + \beta^{SM} M_{it} + \gamma_t^{SX} X_{it}^S, \tag{6}$$

where X_i^S is a vector of controls including children's gender, cohort, age in months, and the index for the socioeconomic status (SES) of the household. The regression also controls for the neighbourhood characteristics that proxy for potential differences in the supply of therapy services. These characteristics are computed based on Census data and include the percentages of children aged 0-4 and 5-9 in the population, percentages of persons with Aboriginal origins, speaking English at home or born in Australia. I control for the neighbourhood's SES using the Index of Relative Socioeconomic Advantage and Disadvantage, which is computed by the Australian Bureau of Statistics.[29] It accounts for a broad range of neighbourhood variables reflecting people's access to material and social resources, and their ability to participate in society (Statistics, 2011).

Columns (1) and (4) of Table 6 show the estimates of Equation (6) for therapy directed at non-cognitive and cognitive development, respectively. Identification of children's non-cognitive delays by teachers is a key predictor of children taking advantage of both non-cognitive and language-oriented therapy. Teachers' identification of delays in language development is an important predictor of children getting learning support or speech therapy, but not behavioural therapy or psychological assessment. Children whose teachers indicate non-cognitive deficits are almost 7 percentage points more likely to receive behavioural therapy as well as 7 percentage points more likely to receive learning or speech therapy. In addition to the teachers' perceptions, the mothers' identification of children's non-cognitive delays predicts the uptake of both types of therapy. Mothers' concern about their children's cognitive development is the strongest predictor of the child's use of learning support or speech therapy.

To an important degree, the uptake of therapy services is predicted by the perceptions of teachers and mothers about children's development. Since teachers in advantaged neighbourhoods are more likely to indicate children's developmental delays as was shown in Section 4 and that these perceptions are transmitted to parents as was shown in Section 5, the reference

[29]See the summary of neighbourhood characteristics in Appendix A.3.

Table 6: Deficit recognition by teachers and mothers and child therapy

	Behavioral or psych therapy			Learning or speech therapy		
	(1)	(2)	(3)	(4)	(5)	(6)
Teach.: Non-cognitive delay	0.069*	0.067*		0.070*	0.069*	
	(0.012)	(0.012)		(0.018)	(0.018)	
Teach.: Cognitive delay	0.013	0.011		0.145*	0.135*	
	(0.014)	(0.014)		(0.022)	(0.022)	
Moth.: Non-cognitive delay	0.154*	0.152*		0.104*	0.101*	
	(0.025)	(0.025)		(0.028)	(0.028)	
Moth.: concern Cognitive	0.058*	0.056*		0.206*	0.200*	
	(0.019)	(0.019)		(0.031)	(0.030)	
Cognitive score		-0.001	-0.009*		-0.019*	-0.041*
		(0.003)	(0.004)		(0.006)	(0.007)
Non-cognitive score		-0.009*	-0.017*		-0.004	-0.019*
		(0.004)	(0.004)		(0.006)	(0.007)
Neighbourhood non-cognitive score		0.005	0.008*		-0.012*	-0.008
		(0.003)	(0.003)		(0.005)	(0.005)
Neighbourhood cognitive score		-0.001	-0.001		0.013	0.015*
		(0.003)	(0.003)		(0.007)	(0.008)
N	4104	4104	4104	4104	4104	4104

Notes: Linear probability regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index, and neighbourhood characteristics from the Census. Standard errors are clustered at the postcode level. Significance level: * 5%

bias is likely to exasperate the gaps in the therapy uptake between children in advantaged and disadvantaged areas.

There are several channels through which the average level of child development in the neighbourhood can be related to the uptake of therapy services. First, in neighbourhoods with high child development levels, the reference bias in teachers' and mothers' perceptions increases the likelihood that teachers and mothers perceive the child to have deficits, resulting in an increased demand for therapy. The reference bias in teachers' perceptions generates a positive association between the average level of child development in the neighbourhood and the uptake of therapy. Second, even though I control for key neighbourhood characteristics potentially related to the available educational opportunities, differences in therapy uptake can be driven by unobserved differences in the supply of educational and developmental services across neighbourhoods. In this case, the availability of child development resources increases both the average level of child development and the uptake of the therapy through the increased supply, generating a positive association between the average level of neighbourhood development and the uptake of therapy. Finally, children in neighbourhoods with high levels of non-cognitive and cognitive development might also rank high in other dimensions of children's well-being that matter for therapy uptake but are omitted from the regression, for example, expressive language skills. In that case, there would be a lower demand for therapy in neighbourhoods with high levels of measured development and a negative association between the therapy uptake and average development levels.

Under the linearity assumed in Equations (4), (5), and (6), I can examine the contribution of reference bias in perceptions to uptake of therapy. Columns (3) and (6) of Table 6 report estimates of a linear probability regression of therapy uptake on interview measures of child development, the average level of neighbourhood non-cognitive and cognitive development, and the set of controls X_{it}^{S}. The estimated effect of average neighbourhood non-cognitive development on the uptake of the therapy reported in columns (3) and (6) is the reduced form effect from potentially all demand-side and supply-side mechanisms described above. There is a positive net relationship between the average level of non-cognitive development in the neighbourhood and the probability of receiving behavioural therapy, which implies that the role of neighbourhood resources or the reference bias is dominating for behavioural therapy uptake. In contrast, there is a negative net relationship between the average level of non-cognitive development and the probability of receiving learning therapy, which implies that neighbourhood non-cognitive development may be strongly correlated with unobserved factors in children's environment of development that relate to therapy uptake.

Columns (2) and (5) control for the measures of teachers' and mothers' perceptions in addition to the variables included in Columns (3) and (6). Therefore, the estimated effect of the average neighbourhood non-cognitive development on the uptake of therapy reported in Column (2) includes all effects estimated in Column (3) except for the effect of the reference bias in teachers' and mothers' assessments. The decrease in the coefficient on the average level of non-cognitive development between Columns (3) and (2) and Columns (6) and (5) indicates that there is some role for reference bias in the uptake of therapy. While these results are not conclusive and should be interpreted with care, they are consistent with residents of neighbourhoods with high average child development levels having higher developmental expectations for their children and, therefore, requesting more professional help.

7 Teachers' perceptions and family environment

In addition to the uptake of therapy services, mothers who think that their children are falling behind are likely to undertake different parental investment strategies. I explore the relationship between mothers' perceptions about children's non-cognitive delays when children are aged 8-9, and a wide range of family investments including parenting style, frequency of household members engaging in development-promoting activities like reading to or with their child, drawing and playing, parental expectations about their children's future, and tutoring.

To understand the role of mothers' perceptions for family investment choices, I estimate a linear regression where the dependent variable is equal to various measures of family-based investment when children are ages 8-9 (I^F_{it}), and the independent variables include the mothers' identification of the children's non-cognitive developmental delays ($M_{i,t}$)

$$I^F_{i,t} = \beta^{F,M} M_{it} + \beta^{F,X}_t X^F_{it} + \beta^{F,M'} M_{it-1} + \beta^{F,I} I^F_{i,t-1} + \epsilon^F_{i,t}, \tag{7}$$

Table 7: Deficit recognition by parents and family investments for children ages 8-9

	Mother: child more difficult			
Dependent variable:	Coef.	SE	N	R2
A: Parental attitudes				
Mother Warmth Score	-0.228*	(0.033)	6584	0.41
Mother Anger Score	0.616*	(0.049)	6582	0.42
Expected educ: college and higher	-0.094*	(0.023)	6186	0.36
B: Weekly investment (times per week)				
Meet with tutor	0.105*	(0.033)	3570	0.09
Talk about school	-0.043	(0.059)	6586	0.09
Read to child	0.092	(0.118)	6621	0.15
Play outside	0.060	(0.103)	6621	0.09
Tell story	0.026	(0.113)	3690	0.13
Music	-0.233	(0.143)	3690	0.13
Play with toys	-0.134	(0.111)	3690	0.08
Draw a picture	0.090	(0.090)	3690	0.11
C: Weekly investment (hours)				
Mom childcare	1.634	(1.725)	2460	0.07
Dad childcare	-1.463	(0.929)	2082	0.10

Notes: Linear regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index, neighbourhood characteristics from the Census, mothers' depression score, lag of mother deficit recognition, lag of the dependent variable. Standard errors are clustered at the postcode level. Significance level: * 5%.

where X_{it}^F is a vector of controls including children's gender, cohort, age in months, and the index for the socioeconomic status (SES) of the household. Control variables also include mothers' depressive symptoms score and neighbourhood characteristics. The standard errors are clustered at the postcode level. I additionally control for lagged values (measured when children are ages 6-7) of mothers' perceptions M_{it-1} and lagged values of investments I_{it-1} to account for the role of idiosyncratic (mis)perceptions and preferences.

Table 7 shows the estimates of Equation (7). There are two noticeable types of effects of deficit recognition. On the one hand, deficit recognition is associated with more compensatory investment in tutoring for the child similar to the findings of Kinsler and Pavan (2021). Combined with the increased use of therapy services, this implies that parents who identify deficits in children are more likely to reach out for professional help for their children. On the other hand, mothers who recognize children's non-cognitive deficits have different attitudes toward their children. They report using less warm and more angry parenting practices. In addition to the change in parenting, mothers report having lower aspirations about their children's future education. Mothers who think that their children have a temperament deficit are 9.4 percentage points less likely to expect that their children will obtain a university education.[30] [31]

Mothers' misperceptions about children are likely to undermine children's developmental trajectories. Parental attitudes have been shown to be important for children's non-cognitive development (Fiorini and Keane, 2014, Falk, Kosse, Pinger, Schildberg-Hörisch, and Deckers, 2021), therefore, overestimation of delays in children's non-cognitive development can have a detrimental effect on children through the decrease in parenting quality. On the other hand, in the presence of evidence on the effectiveness of school-based social and emotional learning programmes the underestimation of deficits by teachers and parents will reinforce existing gaps in non-cognitive skills across less- and more-advantaged areas.[32]

8 Conclusions

In this paper, I explore the effect of the local environment on teachers' and mothers' perceptions of developmental delays in children. Using data from direct non-cognitive development observations collected by LSAC interviewers, I document that teachers in neighbourhoods where misbehaving children are more dominant are less likely to indicate developmental deficits in both non-cognitive and cognitive skills. This implies that the estimates of inequality in non-cognitive or cognitive skills based on teachers' evaluations significantly underestimate the true gaps in child development.

I show that these misperceptions matter beyond obtaining the correct statistics. Children

[30]The estimates reported in Table 7 may suffer from reverse causality, where shifts in parental investments can change mothers' perceptions. For example, if mothers engage in warmer and more encouraging parenting children can share more of their problems and improve mothers' deficit recognition. Appendix E shows the estimates with the potentially endogenous mothers' perceptions instrumented by the indicator for being contacted by the school in the last 12 months. Contact from school is likely to be exogenous to changes in mothers' investment choices. The effects of perceptions on mothers' anger and educational aspirations remain statistically significant, while the effect of perceptions on tutoring remains consistent in magnitude and sign, but becomes imprecisely estimated.

[31]Appendix B explores the role of deficit recognition by teachers in children's development trajectories. This analysis suggests that the effect of deficit recognition on child development is positive, however, the results should be interpreted with caution.

[32]See Kautz, Heckman, Diris, Ter Weel, and Borghans, 2014 for a review of interventions directed at non-cognitive skills.

whose teachers and parents do not recognize their delays in development are less likely to receive professional help like behavioural or learning therapy, or tutoring. Parental attitudes are also affected, and the overestimation of delays is associated with lower educational aspirations and poor parenting choices.

An important direction for future research is exploring various ways to improve the recognition of deficits by teachers and families. To address bias in the assessment of academic skills, educational authorities in many countries are conducting standardized nationwide testing of numeracy and literacy, for example, the National Assessment of Educational Progress in the U.S. or the National Assessment Program – Literacy and Numeracy in Australia. These programs provide both parents and schools with an opportunity to learn about children's true positions in the distribution of cognitive skills for similar-aged children. However, they are expensive and have been shown to depend on multiple factors beyond cognitive skills, including students' effort and test-taking skills(Heckman and Kautz, 2012). Standardized assessments of non-cognitive skills are likely to be costly for both governments and students.

A more efficient approach may be providing training to teachers. I show that teachers' education is associated with improved deficit recognition. Providing psychologist-regulated training allowed interviewers in the LSAC study to collect measures of behaviour for a population of children. Similarly, providing better information about developmental milestones and details of skill accumulation at every childhood stage may allow teachers to assess their students more objectively. In a similar vein, asking teachers to evaluate students based on direct observations and using clear assessment guidelines and objective scales may ameliorate the reference bias in the teachers' assessments.

References

Orazio Attanasio, Flávio Cunha, and Pamela Jervis. Subjective parental beliefs. their measurement and role. Technical report, National Bureau of Economic Research, 2019.

Orazio Attanasio, Áureo De Paula, and Alessandro Toppeta. The persistence of socio-emotional skills: life cycle and intergenerational evidence. Technical report, National Bureau of Economic Research, 2020.

Orazio Attanasio, Sarah Cattan, and Costas Meghir. Early childhood development, human capital, and poverty. Annual Review of Economics, 14:853–892, 2022.

Australian Government Department of Education. Belonging, Being and Becoming: The Early Years Learning Framework for Australia (V2.0). Australian Government Department of Education for the Ministerial Council, 2009.

Kalyca Baker, Mark Sipthorp, and Ben Edwards. A longitudinal measure of socioeconomic position in LSAC. Australian Institute of Family Studies Canberra, ACT, 2017.

Peter Bergman. Parent-child information frictions and human capital investment: Evidence from a field experiment. Journal of political economy, 129(1):286–322, 2021.

Teodora Boneva and Christopher Rauh. Parental beliefs about returns to educational investments—the later the better? Journal of the European Economic Association, 16(6):1669–1711, 2018.

Elizabeth M Caucutt, Lance Lochner, Joseph Mullins, and Youngmin Park. Child skill production: Accounting for parental and market-based time and goods investments. Technical report, National Bureau of Economic Research, 2020.

Juan Chaparro, Aaron Sojourner, and Matthew J Wiswall. Early childhood care and cognitive development. Technical report, National Bureau of Economic Research, 2020.

Raj Chetty, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? evidence from project star. The Quarterly journal of economics, 126(4):1593–1660, 2011.

Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. Econometrica, 78(3):883–931, 2010.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. Review of Economic Studies, 81(1):137–185, 2014.

Emilia Del Bono, Josh Kinsler, and Ronni Pavan. Skill formation and the trouble with child noncognitive skill measures. Technical report, IZA Discussion Papers, 2020.

David J Deming. Four facts about human capital. Journal of Economic Perspectives, 36(3):75–102, 2022.

Rebecca Dizon-Ross. Parents' beliefs about their children's academic ability: Implications for educational investments. American Economic Review, 109(8):2728–2765, 2019.

Christopher Doss, Erin M Fahle, Susanna Loeb, and Benjamin N York. More than just a nudge: Supporting kindergarten parents with differentiated and personalized text messages. Journal of Human Resources, 54(3):567–603, 2019.

Todd Elder and Yuqing Zhou. The black-white gap in noncognitive skills among elementary school children. American Economic Journal: Applied Economics, 13(1):105–132, 2021.

Armin Falk, Fabian Kosse, Pia Pinger, Hannah Schildberg-Hörisch, and Thomas Deckers. Socioeconomic status and inequalities in children's iq and economic preferences. Journal of Political Economy, 129(9):2504–2545, 2021.

Mario Fiorini and Michael P Keane. How the allocation of children's time affects cognitive and noncognitive development. Journal of Labor Economics, 32(4):787–836, 2014.

Jason M Fletcher and Barbara Wolfe. The importance of family income in the formation and evolution of non-cognitive skills in childhood. Economics of education review, 54:143–154, 2016.

Dan D Goldhaber and Dominic J Brewer. Does teacher certification matter? high school teacher certification status and student achievement. Educational evaluation and policy analysis, 22 (2):129–145, 2000.

James J Heckman and Tim Kautz. Hard evidence on soft skills. Labour economics, 19(4):451–464, 2012.

Tim Kautz, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. 2014.

Lukas Kiessling. How do parents perceive the returns to parenting styles and neighborhoods? European Economic Review, 139:103906, 2021.

Josh Kinsler and Ronni Pavan. Local distortions in parental beliefs over child skill. Journal of Political Economy, 129(1):81–100, 2021.

François-Xavier Ladant, Paolo Sestito, and Falco J Bargagli-Stoffi. Rather first in a village than second in rome? the effect of students' class rank in primary school on subsequent academic achievements. arXiv preprint arXiv:2302.10026, 2023.

Matthew Manning, Gabriel TW Wong, Christopher M Fleming, and Susanne Garvis. Is teacher qualification associated with the quality of the early childhood education and care environment? a meta-analytic review. Review of educational research, 89(3):370–415, 2019.

Hong Son Nghiem, Ha Trong Nguyen, Rasheda Khanam, and Luke B Connelly. Does school type affect cognitive and non-cognitive development in children? evidence from australian primary schools. Labour Economics, 33:55–65, 2015.

Cheti Nicoletti and Valentina Tonei. Do parental time investments react to changes in child's skills and health? European Economic Review, 127:103491, 2020.

Richard V Reeves and Dimitrios Halikias. Race gaps in sat scores highlight inequality and hinder upward mobility. Washington, DC: Brookings Institute, 2017.

Australian Bureau Of Statistics. Socio-economic indexes for areas (seifa). Canberra: Australian Bureau of Statistics, 2011.

Gema Zamarro, Collin Hitt, and Ildefonso Mendez. When students don't care: Reexamining international differences in achievement and student effort. Journal of Human Capital, 13(4): 519–552, 2019.

A Additional data description

A.1 Measures of child development and perceptions

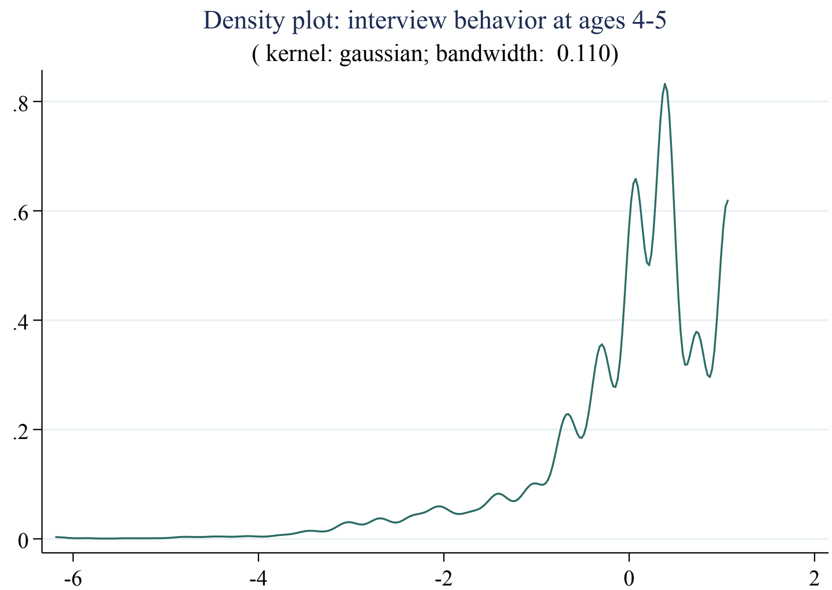


Figure A.1: Distribution of non-cognitive score at ages 4-5

Table A.1: Measures of children's development and perceptions about children by their age

	Ages 4-5			Ages 8-9		
	N	Mean	SD	N	Mean	SD
Teacher: receptive language delay	6598	0.15	0.36			
Teacher: socio-emotional delay	6604	0.21	0.41			
Mother: socio-emotional delay	7144	0.07	0.25	7788	0.07	0.26
Mother: receptive language concern	9365	0.08	0.27			
PPVT score	8672	64.70	6.16	8288	78.72	4.91
WAI score	9077	64.70	8.33			
Socio-emotional score	9189	-0.00	1.00	8252	-0.00	1.00
Average postcode PPVT	8015	-0.00	1.00	5911	0.00	1.00
Average postcode socio-emotional score	8015	-0.00	1.00	5911	0.00	1.00
School contacted about behavior				8304	0.11	0.31

A.2 Measures of background variables

Table A.2: Background variables by children's age

	Ages 4-5			Ages 8-9		
	N	Mean	SD	N	Mean	SD
Child behav. therapy or psych. assessment	9002	0.04	0.21			
Child speech or learning therapy	9002	0.14	0.35			
Share of Kindergarten cohort	9370	0.47	0.50	8417	0.49	0.50
Female child	9370	0.49	0.50	8417	0.49	0.50
Age in months	9370	57.25	2.77	8417	106.10	3.35
Index of SES	9333	0.00	1.00	8349	0.00	1.00
Mother depression	7914	-0.00	1.00	7733	0.00	1.00
Index: mothers' interaction with school				6964	0.00	1.00

A.3 Neighbourhood characteristics

This section provides an additional summary of variables used throughout the analysis. Table A.3 summarizes the neighbourhood characteristics computed by the LSAC. They are based on the statistics from the Census and matched to the data based on the household location. The Index of Neighbourhood Relative Advantage and Disadvantage is computed by the Australian Bureau of Statistics and represents a weighted average of multiple characteristics related to income, employment, education and housing (see Statistics, 2011 for the details of construction). I normalize it by the age group. I additionally control for the age composition of the population and characteristics related to the ethnic or language composition of the population since these variables were not included in the index construction, but can be related to the availability of child development services and language development in the neighbourhood.

Table A.3: Summary of neighbourhood characteristics based on the Census

	Ages 4-5			Ages 8-9		
	N	Mean	SD	N	Mean	SD
Share population age 0-4	9370	6.68	1.37	8416	5.50	1.67
Share population age 5-9	9370	7.12	1.53	8416	4.18	3.18
Index of Neighbourhood Relative Advantage and Disadvantage	9370	0.01	1.00	8409	0.00	1.00
Share population Aboriginal	9370	2.20	3.99	8416	2.33	5.07
Share population English - first language	9370	87.03	13.05	8416	83.93	15.03
Share population born in Australia	9370	86.87	12.41	8416	86.57	14.04

A.4 Family investments

Table A.4 summarizes the family investment measured when children are ages 8-9. Measures of family investment used in the analysis describe parenting, household members' weekly activities with children, the total time mothers and fathers spend with children, and mothers' expectations about children's future educational attainment.

The measures of mothers' parenting include warmth and anger or hostility dimensions. Both measures use the scores constructed by LSAC. The maternal warmth score represents the average for a battery of questions self-reported by mothers measuring the extent to which mothers display warm, affectionate behaviour towards the child. Similarly, the maternal anger score represents the average for a battery of questions measuring the extent to which maternal interactions with children involve disapproval, anger, and the lack of praise. The parenting measures are age-standardized.

The measure of the frequency of tutoring allows mothers' to choose between several categories, including no tutoring, less than once a week, once a week, and more than once a week. I set the weekly number of tutoring sessions to zero if mothers report no tutoring, to 0.5 if mothers report that children meet with a tutor less than once a week, to 1 if mothers report that children have weekly tutoring sessions, and to 1.5 if mothers report that children meet with a tutor more than once a week. On average, children have 0.16 weekly tutoring sessions when they are 8-9 years old.

Mothers' educational aspirations for their children are recovered from the question "Looking ahead, how far do you think study child will go in his/her education?". I create a binary variable equal to one if mothers expect that their child will obtain a university degree or postgraduate qualifications at a university. Almost 70 percent of mothers expect their children to obtain a college degree.

Several variables summarize the frequency with which members of the household engage in educational activities with the child. Mothers are asked to report whether members of the household have engaged in several activities with the child over the past week, including reading, drawing pictures, playing music, singing songs, dancing or doing other musical activities, playing with toys or games indoors, playing outdoors, and telling a story. I transform categorical answers to the weekly frequency as follows: zero if mothers choose ”Not in the past week”, 1.5 times for ”1 or 2 days”, 4 times for ”3-5 days”, and 6.5 times for ”6-7 days”. On average, household members read to their children 2.18 times per week and played outdoors 2.43 times per week. The least frequent activity is drawing with the child, which household members did 1.12 times per week. Mothers reported spending, on average, 26 hours per week actively doing things with their children, while fathers reported spending only 12 hours[33].

Table A.4: Summary of family investment measures

	Ages 8-9		
	N	Mean	SD
Mother warmth	7688	-0.00	1.00
Mother anger	7684	-0.00	1.00
Tutor times per week	4058	0.16	0.41
Mother expects child coll+	8095	0.67	0.47
Weekly times talk school	8369	6.70	0.94
Weekly times read	8378	2.18	2.33
Weekly times play outdoors	8378	2.43	1.95
Weekly times tell story	4048	1.49	1.72
Weekly time music	4048	2.31	2.04
Weekly times play	4048	1.46	1.64
Weekly times draw	4048	1.12	1.41
Mother average weekly time with child	8215	26.17	17.69
Father average weekly time with child	5170	11.97	9.77

A.5 Additional controls

Table A.5 describes additional controls used in Columns (3) and (4) or Table 2 to account for potential idiosyncratic shocks to children's development measures when children are 4-5 years old. The sleeping problems score uses age-standardized answers of mothers to the question about the frequency with which children had problems sleeping over the past month evaluated using the Likert scale.

The ages of the oldest and the youngest children in the children's group are retrieved from the teachers' questionnaire. On average, the youngest child was 50 months old and the oldest

[33] The exact question asked mothers and fathers ”How much time per week do you spend actively doing things with your children, (for example, playing with them, helping them with personal care, teaching, coaching or actively supervising them, getting them to childcare, school or other activities?”

child was almost 67 months old.

During the PPVT test, the majority or parents remained at a distance from their children or observed them. Around 10 percent of parents have actively encouraged their children. Similarly, the majority of children have performed the test with other children not present in the room or remaining at a distance.

Table A.5: Additional control measures

	Ages 4-5		
	N	Mean	SD
Sibling not present	4329	0.62	0.49
Parent not present	4329	0.12	0.32
Interview December	9370	0.00	0.04
Sleeping problems score	8005	0.00	1.00
Parent at a distance	4329	0.42	0.49
Parent observed	4329	0.35	0.48
Parent encouraged	4329	0.10	0.30
Parent interfered	4329	0.01	0.12
Sibling at a distance	4329	0.22	0.41
Sibling observed	4329	0.14	0.35
Sibling encouraged	4329	0.01	0.11
Sibling interfered	4329	0.01	0.09
Interview January	9370	0.00	0.04
Interview February	9370	0.00	0.01
Interview March	9370	0.06	0.23
Interview April	9370	0.12	0.33
Interview May	9370	0.19	0.39
Interview June	9370	0.17	0.38
Interview July	9370	0.20	0.40
Interview August	9370	0.16	0.37
Interview September	9370	0.06	0.24
Interview October	9370	0.02	0.15
Interview November	9370	0.01	0.11

A.6 Teachers' characteristics

Table A.6 summarizes teacher's characteristics when children are ages 4-5. 60 percent of teachers have a bachelor's degree, and 20 percent of teachers have a postgraduate degree. When children are ages 4-5, . The average teaching experience is 15.6 years. The majority of teachers have teaching experience exceeding 10 years. Only 5 percent of teachers have 3 or less years of experience.

Table A.6: Teachers' characteristics for children ages 4-5

	Ages 4-5		
	N	Mean	SD
Child attends daycare	8907	0.26	0.44
teacher_degree	6336	0.64	0.48
Teaching experience 0-5 years	9370	0.11	0.31
Teaching experience 6-10 years	9370	0.15	0.36
Age of youngest child in class	6223	49.13	26.93
Age of oldest child in class	6217	66.67	31.43
Children to qualified staff ratio	6523	14.46	7.02

A.7 Childcare arrangement

Table A.7 summarizes childcare arrangements when children are ages 4-5. The majority of children attend preschool or kindergarten, and 26 percent of children attend daycare. Only a few kids in the sample go to school at this stage.

Table A.7: Childcare arrangements for children ages 4-5

	Ages 4-5		
	N	Mean	SD
Child attends Grade 1	3867	0.00	0.04
Child attends daycare	8907	0.26	0.44
Child attends preschool or kindergarten	8907	0.74	0.44

B The role of deficit recognition for child development

An important question is whether teachers recognizing child developmental delays actually improves children's developmental trajectories. On the one hand, higher uptake of therapy and compensatory investments by schools and families might accelerate child development. On the other hand, the recognition of deficits by teachers might have a negative relative rank effect on children, damaging their self-confidence and later outcomes (Ladant, Sestito, and Bargagli-Stoffi, 2023), and the recognition of deficits by parents might alter their parenting and aspirations in ways that are detrimental to the child development (see Section **??**).

I estimate a value-added regression of the children's measured development at age 8-9 as a function of the teachers' deficit recognition at age 4-5 controlling for the children's development level at age 4-5. The dependent variable is the children's non-cognitive score and the children's PPVT score at ages 8-9.

$$S_{i,t+1} = \beta^T T_{i,t} + \beta^S S_{i,t} + \gamma X_{i,t}^S + \epsilon_i,$$

where control variables include children's gender, cohort, and age in months, household socioeconomic status (SES) index, neighbourhood characteristics from Census, and mothers' depression score.

Table B.1 shows the estimates of the value-added regression (B.1). Columns (1) and (3) show the results of the OLS regression. The coefficient in front of teacher deficit recognition is significant and negative. It can be negative because the recognition of deficits by children actually has a detrimental effect on child development, for example, if the associated stigma damages children's self-esteem and motivation to learn. Alternatively, the negative sign can be driven by the unobserved heterogeneity with teachers deciding that the child has developmental delays due to unobserved negative factors related to the children's development or family environment that affect both the teachers' evaluations and the children's future outcomes.

Columns (2) and (4) of Table B.1 address the endogeneity issue by instrumenting for the deficit recognition by teachers with the average neighbourhood child development levels. It can be seen that addressing the endogeneity problem makes the estimated coefficient in front of the deficit recognition by teachers positive. This implies that the role of unobserved factors in affecting both teachers' deficit recognition and child development is strong. Teachers' deficit recognition seems to be advantageous and particularly important for the development of non-cognitive skills.

These estimates should be taken with some scrutiny since without the additional information about the peer formation process, positive peer effects of being surrounded by children with better skill levels might confound the estimates reported in Columns (2) and (4).

Table B.1: Teachers' perceptions and child development

	Non-cognitive score at 8-9		Cognitive score at 8-9	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	b/se	b/se	b/se	b/se
Non-cognitive score	0.09*	0.20*	0.02	0.06
	(0.02)	(0.05)	(0.01)	(0.04)
Cognitive score	0.04*	0.20*	0.39*	0.44*
	(0.02)	(0.07)	(0.01)	(0.06)
Teach: emot or lang delay	-0.08*	1.99*	-0.13*	0.72
	(0.04)	(0.89)	(0.03)	(0.70)
N	5448	4610	5466	4625
F stat.		9.00		8.62

Notes: Columns (1) and (3) report OLS estimates. Columns (2) and (4) report TSLS estimates: teacher-identified deficit instrumented by postcode average development levels. Control: control variables include children's gender, cohort, and age in months, household socioeconomic status (SES) index, and neighbourhood characteristics from Census. Significance level: * 5%

C Teachers' quality and deficit recognition in children

This section performs the sensitivity analysis for the estimated role of teachers' qualifications in the recognition of developmental deficits in children. Column (1) reports the estimates of a regression similar to the one reported in Column (1) of Table 3, but for the subsample of children in the first quartile of non-cognitive scores. The role of teachers' experience for deficit recognition remains insignificant. Column (2) reports the estimates for a subsample of children with the interview measure of non-cognitive development below the median, similarly to Column (1) of Table 3. Instead of measuring teachers' experience in years, I create indicators for teachers' experience being between 0 and 3 years, between 4 and 10 years, or over 10 years. The estimated coefficients of experience dummies measure the difference in the likelihood of deficit recognition compared to teachers with more than 10 years of experience. The teachers' experience is not a significant predictor of deficit recognition. Finally, Column (3) reports the estimates for children who have demonstrated negative behaviour or the lack of focus during the interview (see Figure 1). The role of teachers' experience for deficit recognition remains very limited, while the role of education remains significant and robust to how I define the subsample of children with non-cognitive deficits.

Table C.1: Teachers' qualifications and non-cognitive deficits recognition in children with deficits

	(1) Int behavior first qrt b/se	(2) Int behavior low b/se	(3) Negative/unfocused b/se
Average postcode behavior	0.027	0.066**	0.026
during int	(0.044)	(0.026)	(0.034)
PPVT	-0.079***	-0.064***	-0.072***
	(0.015)	(0.010)	(0.013)
Average postcode PPVT	0.059*	0.043**	0.068**
	(0.033)	(0.022)	(0.029)
Teacher bachelor	0.096***	0.062***	0.064***
	(0.028)	(0.019)	(0.024)
Teacher postgrad	0.081**	0.058**	0.061*
	(0.037)	(0.025)	(0.033)
Teaching experience	0.002		
	(0.001)		
Behavior during int		-0.043***	
		(0.010)	
Teaching experience 0-3		-0.026	-0.019
years		(0.030)	(0.040)
Teaching experience 4-10		-0.009	-0.039*
years		(0.019)	(0.024)
N	1287	2528	1768

Notes: Linear probability regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. SE clustered at the postcode level. Column (1) defines children with non-cognitive deficits if the measured score is in the first quartile of non-cognitive scores. Column (2) defines children with non-cognitive deficits if the measured non-cognitive score is below the median of non-cognitive scores. Reference (omitted) group: teachers with 11+ years of experience. Column (3) defines children with non-cognitive deficits if children have shown negative behaviour or were not focused during the cognitive test. Reference (omitted) group: teachers with 11+ years of experience. Significance level: * 5%

D Measured development and later outcomes

I test whether interview measures of a children's non-cognitive and cognitive development are significant predictors of child outcomes in the later periods. Better child non-cognitive scores at ages 4-5 is associated with a lower probability of repeating the grade by ages 12-13 and higher scores for the nationwide tests of Grade 9 reading and numeracy.

Table D.2: Interview development measures at age 4-5 and later child outcomes.

	Repeated grade by ages 12-13	Grade 9 Reading	Grade 9 Math
Non-cognitive score	-0.010*	4.199*	4.440*
	(0.003)	(0.873)	(0.929)
Cognitive score score	-0.011*	17.490*	12.168*
	(0.003)	(0.869)	(0.920)
N	6699	5739	5678

Notes: Linear regressions. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. SE clustered at the postcode level.

E Addressing reverse causality in mothers' perceptions and investment

This section addresses the potential endogeneity problem in the estimation of the impact of mothers' perceptions about children's developmental deficits on investment choices reported in Table 7. It is plausible that some unobserved shocks resulted in changes to the family investment and contemporaneous updating of mothers' perceptions about their children. For example, a negative shock to the family environment or health not captured by the set of controls can result in more angry parenting toward the child and an increased likelihood of perceiving the child as "difficult". It is also possible that the estimation suffers from a reverse causality problem. For example, more time spent with the child can make mothers increasingly aware of their children's non-cognitive deficits.

To address these problems I estimate a TSLS regression of the Equation (7) instrumenting for mothers' deficit recognition with an indicator for being contacted by the school about child behaviour within the last 12 months. Contact from school is likely to be exogenous to the family environment after conditioning on the set of controls and lags of perceptions and investment. It is also a significant predictor of mothers' perceptions as discussed in Subsection 5.

Table E.1 reports the results for parental attitudes and tutoring. It can be seen that the effect of non-cognitive delay recognition by mothers on anger and educational aspirations remains statistically significant. Mothers who recognize non-cognitive delays in their children tend to use more angry parenting practices and have lower educational aspirations for their children. The effect of mothers' deficit recognition on tutoring remains positive but becomes imprecisely estimated.

Table E.1: TSLS estimation of the role of mothers' perceptions for the investment choices

	Warmth	Anger	Tutor	Exp coll+
Mother: Non-cognitive	0.113	2.152*	0.065	-0.850*
delay	(0.279)	(0.354)	(0.171)	(0.185)
N	6556	6554	3570	6186
F stat.	77.24	66.49	50.14	65.21

Notes: TSLS regression: mothers' deficit recognition instrumented with indicator: being contacted by school about children's behaviour in the last 12 months. Control: demographic variables, mothers' depression score, neighbourhood characteristics, lag of mother deficit recognition, lag of the dependent variable. SE clustered at the postcode level. Significance levels: *** 1% ** 5% * 10%.

F Robustness to the construction of average neighbourhood development

This section explores the sensitivity of the reference bias estimates to the choice of controls for the SES and details of the construction of the neighbourhood's average child development.

I show that estimates remain consistent if I construct the average for postcodes with a different minimum number of children with measured development per postcode (compared to 10+ measurements used for the baseline estimates), or if I construct the neighbourhood average using only measures of children from the same cohort (similar age, same year), and using the measures for children from both cohorts and all ages adjusted for the cohort and age effect (all ages, all years). A broader measure allows for decreasing the measurement error of the neighbourhood's average development level at the cost of introducing assumptions about the change in the distribution of development levels across time or age.

Columns (1) and (2) present the estimates of the linear probability model specified by Equation (4), including the full set of controls for the family's socio-economic status instead of one index. The control variables include mothers' age, number of siblings, family income, mothers' education, mothers' marital status, whether English is the main language spoken at home, and mothers' employment status in addition to children's cohort, gender, and age in months.

Columns (3) and (4) report the estimates with the postcode average constructed in a similar way as the one that is used for main results (using measures for children of the same age and both cohorts) for postcodes where more than 15 measures are reported. Columns (5) and (6) report the estimates with the postcode average constructed in a similar way as the one that is used for main results (using measures for children of the same age and both cohorts) for postcodes where more than 5 measures are reported. Columns (7) and (8) present estimates received using the average postcode measure constructed for the children of the same age during the same year (same cohort) if more than 5 measures are available in a given postcode. Columns (9) and (10) document the estimates with the postcode average constructed from the measures for all available ages and years after first running a regression of measures on the cohort and age dummies and then using the residuals from this regression to construct the postcode averages are more than 15 measures are available for a given postcode. It can be seen that all of these modifications to the estimation procedure produce estimates corroborating the role of the children's own non-cognitive development and the positive association between the postcode average development level and the probability that teachers recognize children's developmental delays.

Table F.1: Robustness table: bias in teachers' perceptions

	Full SES controls		15+ obs per postcode		5+ obs per postcode		Same age/year		All age/year	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Emot	Rec. lang.	Emot	Rec. lang.	Emot	Rec. lang.	Emot	Rec. lang.	Emot	Rec. lang.
Average postcode PPVT	0.025	0.042***	0.025	0.039**	0.032**	0.043***	0.025*	0.037***	0.037**	0.042**
	(0.016)	(0.014)	(0.018)	(0.016)	(0.016)	(0.014)	(0.013)	(0.011)	(0.017)	(0.016)
PPVT score	-0.052***	-0.082***	-0.052***	-0.082***	-0.051***	-0.080***	-0.052***	-0.080***	-0.051***	-0.080***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Average postcode behavior score	0.066***	0.045***	0.064***	0.050***	0.069***	0.050***	0.034**	0.025**	0.079***	0.037*
	(0.019)	(0.016)	(0.020)	(0.017)	(0.019)	(0.016)	(0.014)	(0.013)	(0.023)	(0.020)
Int. behavior score	-0.043***	-0.030***	-0.043***	-0.033***	-0.043***	-0.030***	-0.041***	-0.032***	-0.042***	-0.029***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
N	5258	5254	4929	4924	5387	5383	5237	5233	5315	5311

G Measurement error in average neighbourhood development

This section explores the role of the measurement error in children's individual development and in the constructed average neighbourhood development levels. Columns (1) and (4) duplicate the baseline estimates of linear probability regression of teachers' deficit recognition reported in Columns (2) and (4) of Table 1. Columns (2) and (5) duplicate the results reported in Columns (5) and (6) of Table 2. They report the results of a linear TSLS regression with children's focus during the cognitive test instrumented by degrees of negative and positive response, and children's PPVT score instrumented by WAI score. Finally, Columns (3) and (6) report results adjusted for the potential measurement error in both children's individual development measures and in neighbourhood average development measures. In addition to instrumenting for the measures of individual child development, I instrument for the postcode average level of focus during the cognitive test with average degrees of positive and negative response during the interview, and for the postcode average PPVT with the postcode average WAI score. Adjusting for the potential measurement error in neighbourhood development measures further reinforces the positive and significant role of the average postcode non-cognitive score on the probability that the teacher will identify deficits in both non-cognitive and cognitive dimensions of child development.

Table G.1: Teachers' perceptions and average neighbourhood development

	Non-cognitive delay			Cognitive delay		
	(1)	(2)	(3)	(4)	(5)	(6)
	Logit Avg. Marg. Effect	Extra control	Meas. error adj.	Logit Avg. Marg. Effect	Extra control	Meas. error adj.
Neighbourhood cognitive	0.010	0.003	-0.064	0.016*	0.030*	-0.003
score	(0.006)	(0.018)	(0.045)	(0.005)	(0.013)	(0.036)
Cognitive score score	-0.052*	-0.045	-0.103	-0.082*	-0.155*	-0.197*
	(0.007)	(0.067)	(0.054)	(0.006)	(0.052)	(0.045)
Neighbourhood	0.020*	0.050*	0.162*	0.014*	0.025*	0.097*
non-cognitive score	(0.006)	(0.013)	(0.059)	(0.005)	(0.010)	(0.046)
Non-cognitive score	-0.043*	-0.386*	-0.283*	-0.030*	-0.161	-0.090
	(0.007)	(0.112)	(0.086)	(0.006)	(0.087)	(0.077)
N	5258	5215	5215	5254	5211	5211

H Sensitivity to childcare arrangement

Since the childcare arrangement might play an important role in the therapy uptake and teachers' deficit recognition, this section explores the sensitivity of results to controlling for the arrangement. Columns (2) and (5) in Table 2 show estimates of regression specifications that control for whether children attend childcare. The reference bias persists after controlling for the childcare arrangement. Table H.1 shows the estimates reported in Table 3 with the added control for whether children attend daycare versus preschool or kindergarten. Teachers in daycare institutions are less likely to report delays in child development. However, the effect of teacher's education on deficit recognition is preserved conditional on the childcare arrangement.

Table H.1: Teacher quality and deficit recognition

	Non-cognitive delay		Cognitive delay	
	(1)	(2)	(3)	(4)
	Non-cogn. score low	Non-cogn. score high	Cognitive score low	Cognitive s
Non-cognitive score	-0.043*	0.011	-0.048*	-0.01
	(0.010)	(0.024)	(0.009)	(0.00
Neighbourhood	0.020*	0.017*	0.022*	0.00
non-cognitive score	(0.008)	(0.008)	(0.008)	(0.00
Cognitive score score	-0.064*	-0.039*	-0.115*	-0.04
	(0.011)	(0.009)	(0.014)	(0.01
Neighbourhood cognitive	0.017*	0.002	0.017*	0.014
score	(0.008)	(0.008)	(0.009)	(0.00
Teacher coll+	0.041*	0.011	0.048*	0.00
	(0.019)	(0.017)	(0.016)	(0.01
Teaching experience	-0.001	0.000	0.000	-0.00
	(0.001)	(0.001)	(0.001)	(0.00
Child attends daycare	-0.050*	-0.035	-0.054*	-0.01
	(0.024)	(0.020)	(0.020)	(0.01
N	2449	2382	2482	234

Notes: Linear probability regression for high-development (measured development higher than the median) and low-development (measured development lower than the median) samples. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. Standard errors are clustered at the postcode level. Significance level: * 5%.

Table H.2 shows the estimates reported in Table 6 with the added control for whether children attend daycare versus preschool or kindergarten. Attending daycare decreases the likelihood of children receiving learning or speech therapy. However, the effect of teachers' and mothers' perceptions on the uptake of therapy is preserved.

Table H.2: Perceptions and therapy uptake

	Behavioral or psych therapy	Learning or speech therapy
	(1)	(2)
Teach.: emot deficit	0.067*	0.067*
	(0.012)	(0.018)
Teach.: Cognitive delay	0.011	0.134*
	(0.014)	(0.022)
Moth.:child temperament deficit	0.152*	0.103*
	(0.025)	(0.028)
Moth.: concern receptive lang	0.056*	0.200*
	(0.019)	(0.030)
Cognitive score score	-0.001	-0.020*
	(0.003)	(0.006)
Non-cognitive score	-0.009*	-0.004
	(0.004)	(0.006)
Neighbourhood non-cognitive score	0.005	-0.011*
	(0.003)	(0.005)
Neighbourhood cognitive score	-0.001	0.012
	(0.003)	(0.007)
Child attends daycare	-0.002	-0.034*
	(0.007)	(0.013)
N	4104	4104

Notes: Linear probability regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index, and neighbourhood characteristics from the Census. Standard errors are clustered at the postcode level. Significance level: * 5%.

I Sensitivity to the selected measure of teachers' perceptions

This section estimates the reference bias in teachers' perceptions about children's non-cognitive skills using an alternative measure of teacher's perceptions. I construct a continuous measure of teachers' assessment of children's non-cognitive problems based on a selection of subquestions from the Strength and Difficulty Questionnaire. The Strength and Difficulty Questionnaire is a battery of questions commonly used to evaluate children's non-cognitive development (Fiorini and Keane, 2014, Nicoletti and Tonei, 2020). I select subquestions that relate to behaviours that were evaluated during the interview.

Table I.3 illustrates the results of the regression specified in (4) where the dependent variable is the constructed index of children's non-cognitive problems for children ages 4-5 and 8-9. The coefficient of the average postcode non-cognitive development is significant and positive and does not decrease with children's age.

Table I.3: Reference bias in alternative measures of teachers' perception about noncognitive skills

	Ages 4-5 (1)	Ages 8-9 (2)
Neighbourhood cognitive score	-0.005	0.002
	(0.017)	(0.020)
Cognitive score score	-0.076*	-0.051*
	(0.015)	(0.014)
Neighbourhood non-cognitive score	0.036*	0.037*
	(0.017)	(0.017)
Non-cognitive score	-0.070*	-0.054*
	(0.016)	(0.016)
N	5055	4679

Notes: Linear regression. Control: children's gender, cohort, and age in months, household socioeconomic status (SES) index. Standard errors are clustered at the postcode level. Significance level: * 5%.