Mathematics Trajectories and Risk Factors During Childhood

LEBIHAN, Laetitia and MAO TAKONGMO, Charles Olivier

University of Ottawa, University of Ottawa

2018

Mathematics Trajectories and Risk Factors During Childhood*

Laetitia LEBIHAN† Charles Olivier MAO TAKONGMO‡

August 2018

Abstract

In this paper, we use the National Longitudinal Survey of Children and Youth to identify the trajectory of mathematical abilities among Canadian children 7 to 15 years old. We also analyse families and personal characteristics during early childhood that may influence the likelihood of being in one of these abilities groups. We identify three trajectory groups: average abilities (47.6%), high abilities (30.1%), and low abilities (22.3%). Our results also show that maternal education is one of the most important predictors for a low mathematics abilities trajectory. Cognitive score at ages 4 to 5 is also a good indicator of future academic success. Finally, children at risk are those whose parents have low parenting skills.

Key words: group-based trajectory modelling, mathematical abilities, risk factors, Canadian children.

*This paper previously circulated under the title "Academic Achievement Trajectories and Risk Factors During Early Childhood." We wish to thank professors Steven Ambler, Jean-Marie Dufour, Catherine Haeck, Pierre Lefebvre, Philip Merrigan, and Victoria Zinde-Walsh. The analysis is based on Statistics Canada's National Longitudinal Survey of Children and Youth restricted-access Micro Data Files.

†Part-time Professor, University of Ottawa. Email: llebihan@uottawa.ca. Website: https://sites.google.com/site/laetitiamlebihan.

‡Assistant Professor, University of Ottawa, 120 University, Social Sciences Building, Room 9005, Ottawa, Ontario, Canada, K1N 6N5; e-mail: cmaotako@uottawa.ca Web page: https://sites.google.com/site/maotakongmocharles/

1 Introduction

Evidence indicates the importance of skills in early childhood for academic success and future outcomes, such as employment and wages (Card, 1999; Green & Riddell, 2003; Heckman, 2006). In particular, mathematics skills have greater long-term positive effects on future wages compared to other cognitive outcomes, such as reading or vocabulary (Rivera-Batiz, 1992; Murnane et al., 1995; Rose & Betts, 2004). Similarly, adults who are proficient in reading are still more likely to be unemployed (and/or less likely to be promoted if they are employed) if they are not qualified in mathematics (Bynner & Parsons, 2006). Several studies of early intervention programs also demonstrate the importance of the child's environment, based on family characteristics and the personal characteristics of the child (e.g., behaviour, adaptive capacities), in predicting future outcomes. For these reasons, the abilities that a child possesses at the beginning of schooling might result in different achievement behaviours later (Duncan et al., 2007).

Many studies report distinct academic achievement trajectories for children (Caro, 2009; Herbers et al., 2012). An implicit assumption in the identification of heterogeneity in a population is that a group of individuals who follow an atypical developmental trajectory also has specific risk factors. Thus, some child characteristics such as behaviour or cognitive abilities when entering school may have a negative effect on mathematics achievement and development in general (Duncan et al., 2007). Children living in low-income households (Jordan et al., 2006) or who have a young mother at birth (Corcoran, 1998; Dahinten et al., 2007) may also have lower math scores. Similarly, living in a family environment of poor quality may also affect the child's cognitive development (Todd & Wolpin, 2003, 2007).

The group-based trajectory approach proposed by D. S. Nagin (2005) has become very popular in the literature. The method provides a way to identify clusters of individuals who follow a similar progression in some variable over age or time and to determine the predictors of membership in each trajectory group. Several studies have used this method to analyse the cognitive developmental trajectories of children or youth; they then linked those trajectories to socio-economic or personal characteristics of the children in the earlier stage of their development. For instance, M. H. Gagné et al. (2018) use this group-based approach to analyse the mathematics and language trajectories of adolescents aged 14 to 19 and identify the socio-economic factors that predict such variations. However, their study was restricted to the province of British Columbia in Canada. Garon-Carrier et al. (2018) use the same method to assess the developmental trajectories of number knowledge and mathematics abilities from late preschool to school entry and elementary school in Quebec. They also associate each trajectory to children's socio-economic conditions. Sutcliffe et al. (2017) use this approach to analyse the impact that early entry to care has on trajectory group membership in children's educational progress in England. Iruka et al. (2018) use the same method to explore the factors that contribute to academic and socio-emotional competencies of young children in North Carolina and Pennsylvania.

To the best of our knowledge, this is the first study to use group-based trajectory analysis for mathematics scores with representative data for the whole Canadian population. Previous studies focused on only one subgroup or region (M. H. Gagné et al., 2018; Garon-Carrier et al., 2018). Moreover, the majority of Canadian studies on mathematics achievement trajectories focus on a specific risk factor, such as the effect of early motherhood (Dahinten et al., 2007) or parental socio-economic status (Caro, 2009). The present study identifies the mathematics achievement trajectories for children aged 7 to 15 years and the risk factors associated with low mathematics skills trajectories during early childhood. We use a greater number of risk factors from early childhood and study the cognitive scores and behaviour of the child and the child's family characteristics.

It has been shown that certain periods are more critical than others in a child's cognitive development: acting as soon as possible, particularly during early childhood, is essential to limit factors that may be harmful to the child (Shore, 1997). The ability to identify different groups of individuals and the risk factors associated with each group is essential in understanding how to respond politically. We can then target public interventions to specific groups so that children can advance in their mathematics capabilities, ultimately reducing potential gaps among individuals. Indeed, several experimental studies show that high-quality programs for at-risk preschool children yielded cognitive and academic gains and reduced behavioural problems (Karoly et al., 2006). These early intervention programs also have long-term benefits, such as reductions in the rates of grade retention and drop-outs (Reynolds & Temple, 1998; Campbell et al., 2002).

Unlike previous studies, this paper uses representative data for the whole population of Canada. We use data from the National Longitudinal Survey of Children and Youth (NLSCY) (8 biennial waves: 1994 to 2009), which constitute a representative sample of the Canadian population, to identify the mathematical abilities trajectories of Canadian children 7 to 15 years old. We also analyse family and personal characteristics during early childhood that may influence the likelihood of being sorted into each abilities group. The following family and personal characteristics during early childhood are considered: i) child characteristics (sex, behaviour, and cognitive score); ii) family characteristics (low-educated mother, insufficient household income, single parent, early motherhood, presence of siblings in the household); and iii) parenting practices. Mathematics achievements are measured when the child is between 7 and 15 years old and risk factors are measured at the age of 4 to 7 years.

We identify three trajectory groups: average abilities (47.6%), high abilities (30.1%), and low abilities (22.3%). Our results also show that maternal education is one of the most important predictors for the low mathematics abilities trajectory. Thus, a child with a low-educated mother is more likely to be in the low skills group than in the average group. Cognitive score at ages 4 to 5 is also a good indicator of future academic success: A low score is a good indicator of students who might struggle in early adolescence. Children at risk are also those whose parents have low parenting skills.

This paper is structured as follows. Section 2 describes the methodology. The data set used is presented in Section 3. Empirical results are presented in Section 4 and discussed in Section 5. Section 6 concludes the paper.

2 Methodology

This section presents a brief summary of the group-based trajectory modelling from D. S. Nagin (2005). This method provides a flexible and easy way to identify clusters of individuals who follow a similar evolution in a variable of interest over time. The aim of the method is to determine the distribution of the variable of interest conditional to age.

Let $Y_i = (y_{i1}, y_{i2}, ..., y_{iT})$ represent individual $i's$ longitudinal random variable of interest and $Age_i = (a_{i1}, a_{i2}, ..., a_{iT})$ represent the age of individual i at each period of time t. It is assumed that the population distribution of development trajectories comes from a finite mixture of unknown order K (D. S. Nagin, 2005; Jones & Nagin, 2007, 2013).

The likelihood of observing a mathematical measure Y_i for individual i, conditional on the number of groups K, can be written as:

$$P(Y_i) = \sum_{k=1}^{K} \pi_k . P(Y_i \mid k), \tag{1}$$

where π_k is the probability that a randomly chosen population member belongs to group k, and $P(Y_i \mid k)$ is the probability of observing Y_i given the membership of group k.

The model also assumes that, conditional on membership of group k, the random variables y_{it}, $t = 1, 2, ..., T$ are independent. We then have:

$$P(Y_i \mid k) = \prod_{t=1}^{T} p(y_{it} \mid k), \tag{2}$$

where $p(y_{it} \mid k)$ is the probability distribution function of y_{it} given membership in group k at time t.

2.1 The probability distribution function of y_{it} given membership in group k at time t, $p(y_{it} \mid k)$

The choice of $p(y_{it} \mid k)$ depends on the type of variable used. For count variables, $p(y_{it} \mid k)$ can be a Poisson distribution. For binary variables, the binary logit distribution can be used to model $p(y_{it} \mid k)$. Finally, for censored variables–in which there are clusters of data at the scale minimum or maximum– $p(y_{it} \mid k)$ can follow the censored normal distribution. We use the normal censored model in this paper.[1]

[1]The normal censored model is also suitable for continuous data that are normally distributed, with or without censorship. The uncensored case is treated by specifying minimum and maximum values that are outside the range

The normal censored model is used to analyse repeated measures, continuous scales that can be censored by a minimum (y_{min}) or maximum scale (y_{max}) (e.g., longitudinal data with a scale of depression symptoms (D. Nagin & Tremblay, 1999)).

Let $\beta^k Age_{1,2,3} = \beta_{0,k} + \beta_{1,k} age_{it} + \beta_{2,k} age_{it}^2 + \beta_{3,k} age_{it}^3$. In that case, the probability distribution function of y_{it} given membership in group k at time t is written as

$$p(y_{it} = y_{min} \mid k) = \Phi\left(\frac{y_{min} - \beta^k Age_{1,2,3}}{\sigma}\right),$$

$$p(y_{it} \mid k) = \frac{1}{\sigma}\phi\left(\frac{y_{it} - \beta^k Age_{1,2,3}}{\sigma}\right) \text{ for } \quad y_{min} \leq y_{it} \leq y_{max},$$

$$p(y_{it} = y_{max} \mid k) = 1 - \Phi\left(\frac{y_{max} - \beta^k Age_{1,2,3}}{\sigma}\right),$$

where ϕ and Φ are, respectively, the density function and the cumulative distribution function of a normal random variable with mean $\beta^k Age_{1,2,3}$ and standard deviation σ.

2.2 The probability of group membership, π_k, and associated risk factors

The model also helps to analyse the effect of a time-stable covariate (risk factors), z_i, on the probability of group membership, π_k. The effect of risk factors on group membership is modelled by a functional relationship between π_k and the vector of multiple individual characteristics z_i for individual i using a generalised logistic function as follows:

$$\pi_k(z_i) = \frac{e^{z_i \theta_k}}{\sum_{k=1}^{K} e^{z_i \theta_k}} \tag{3}$$

where $\theta_1 = 0$. The parameter θ_k captures the impact of z_i on the probability of membership k, denoted π_k.

2.3 The posterior probability of group membership

The Bayesian theorem is used to compute the posterior probability that individual i has group membership k (D. S. Nagin, 2005). This is the probability that individual i has group membership k given the individual's characteristics Y_i. The posterior probability of group membership k given Y_i can be written as:

$$P(k \mid Y_i) = \frac{P(Y_i \mid k)\pi_k}{\sum_{j=1}^{K} P(Y_i \mid j)\pi_j} \tag{4}$$

of observed data values (Jones et al., 2001; D. S. Nagin, 2005; Jones & Nagin, 2013).

2.4 Practical estimation procedure

The parameter θ_k and the slope parameters, $\beta^k = (\beta_{0,k}, \beta_{1,k}, \beta_{2,k}, \beta_{3,k})$, are estimated simultaneously by maximum likelihood for a varying number of groups, called models. For each model, the Bayesian information criterion (BIC) is calculated:

$$BIC = -2log(L) + m.log(n) \tag{5}$$

where L is the maximised likelihood of the model; n is the sample size; and m is the number of parameters in the model. The chosen model (number of groups K) is the one with the highest BIC value.

The "posterior probability" in equation 4 is calculated, based on equations (2) and (3). For each individual, the estimated model coefficients can be used to calculate the probability that the individual belongs to each group. Individuals are assigned to the group for which their posterior membership probability is larger. Probabilities greater than or equal to 0.7-0.8 are considered a good fit for the group to which each individual is assigned. The estimated proportion of individuals belonging to each trajectory group is then calculated. Group-based trajectory analysis was performed with PROC TRAJ within Stata (D. S. Nagin, 2005; Jones & Nagin, 2013).

3 Data

The NLSCY is a long-term survey designed to provide information about the development and well-being of Canadian children and youths.[2] This survey was conducted every two years beginning in 1994-95 (wave 1) and ending in 2008-09 (wave 8). A cohort of approximately 2,000 children aged 0 to 11 years was selected in the initial cycle and followed longitudinally through the entire survey. Then, in every wave, new cohorts of children aged 0 to 1 year were added and followed until ages 4 to 5. Given the complex sampling design of the NLSCY, all estimates are performed using the sample weights provided by Statistics Canada.

3.1 Measure of mathematical achievement

The CAT/2 mathematical test, the results of which were available in NLSCY data, measures the mathematics skills of school-age children. The CAT/2 test is a shorter version of the Mathematics Computation Test taken from the Canadian Achievement Tests, 2nd edition. The test consists of 20 calculation questions and measures the ability of the child/youth to perform addition, subtraction, multiplication, and division of whole numbers, decimals, fractions, negatives, and exponents.

[2]The target population was restricted to ten Canadian provinces and excluded children living on Aboriginal reserves and with full-time members of the Canadian Armed Forces. These exclusions represent approximately 2% of the Canadian population.

Problems involving percentages and the order of operations are also evaluated. The test is administered to children aged 7 to 15 years enrolled in grades 2 to 10. The test levels (2 to 10) vary with the school grade of the child. For example, grade 2 children (aged 7) were given the level 2 test, grade 3 children (aged 8 years) were given the test 3 level, and so on. All children who passed the math test were awarded a raw score and a standardised score. The raw score is simply the number of correct answers to the test, and the standardised score is calculated according to the standards set in 1992 by the Canadian Test Centre. We use the standardised scores because they represent the mathematical skill level that the child has reached, which allows us to track the child's mathematical progress through the years.

Math scores are used only during the last five waves (waves 4-8). In 1994-95, during the first wave, a high proportion of children obtained perfect scores, making it impossible to distinguish the true top performers. Subsequently, because of this ceiling effect, the difficulty level of the tests was adjusted in 1996-97 (wave 2). Nevertheless, during waves 2 and 3, the test was administered by the teacher, leading to a low response rate (74% in 1996-97 and 54% in 1998-99). Statistics Canada decided that from the year 2000 (wave 4), the math test would be administered at home by the interviewer rather than at school, and almost all eligible students (about 90%) responded. Therefore, we select children between 6 and 9 years old in wave 4 (2000-01) with at least four and up to five math scores.

Table 1 shows the descriptive statistics of the standardised math scores (mean, standard deviation, and number of individuals). These statistics are classified by wave and the age of the child.

3.2 Risk factors

Risk factors are measured during early childhood (0-7 years) and extended over the first three waves of the NLSCY. The characteristics of early childhood are divided into three categories: i) child characteristics, ii) family characteristics, and iii) parental characteristics. Following Côté et al. (2006), all risk factors are binary, with a score of 1 if the risk factor is present and can have a negative impact on the child's cognitive development and 0 otherwise. For continuous scales, unless otherwise specified, a score of 1 is assigned for scores equal to or above the 75th percentile of the distribution and 0 if below.

Child characteristics

First, the sex of the child was coded as 0 for boys and 1 for girls. Second, the child's Peabody Picture Vocabulary Test (PPVT) score is measured. This is a vocabulary test in which the child observes pictures on an easel and identifies the picture that matches the word the interviewer reads out. We use the standardised PPVT-R test to reflect the age of the child, as opposed to the raw score based only on the number of correct answers. Evidence suggests the PPVT to be a good predictor of reading and writing abilities and, consequently, the child's academic success

(Hoddinott et al., 2002). These data also allow us to observe whether early vocabulary skills are related to future math skills. Indeed, several studies show a significant link between the PPVT score and future cognitive outcomes (Romano et al., 2010; Baker, 2011). The PPVT score is standardised, with a mean of 100 and a standard deviation of 15. We code 1 for children with a low PPVT score (less than 85) and 0 otherwise. Because the test is administered when the child is 4 to 5 years old, we use waves 2 and 3 to be able to capture it.

Third, behavioural scores available in the NLSCY measure the social and emotional development of children aged 4 to 11 years based on the frequency of events related to the behaviour of the child. The following scores are used: i) hyperactivity/inattention (score ranging from 0 to 16), ii) physical aggression (score ranging from 0 to 12), and iii) indirect aggression (score ranging from 0 to 10). Evidence shows that non-cognitive skills (social and behavioural skills) are significant in predicting future outcomes, such as employment and wages (Carneiro et al., 2007). It is natural to expect that non-cognitive skills also affect cognitive development. The socio-emotional abilities of the child can affect individual learning and classroom dynamics. Socio-emotional or behavioural problems can generate child-teacher conflicts and lead to social exclusion, which may reduce the child's participation in educational activities and consequently affect academic achievement (Duncan et al., 2007). We code 1 for children with a high behavioural score (at or above the 75th percentile) and 0 otherwise. Variables are measured in wave 3 because this wave concerns only children aged 4 to 11 years.

Family characteristics

First, maternal education was treated as a dummy variable to differentiate between mothers who had a high-school diploma or more (0) and those who did not (1). Maternal education seems to be the most important factor in a child's cognitive development because the knowledge that the mother can transmit depends on her education. Indeed, a child's cognitive skills are promoted by the "quality" of interactions had with his or her mother, and the mother's level of education is a good indicator of this quality (Verstraete, 2006). Many studies highlight the important role that the mother's education and social capabilities play in child development (NICHD, 2002; L. G. Gagné, 2003). This variable is measured in wave 3.

Second, the age of the biological mother at birth has been transformed into a binary variable with a value of 1 if the mother was 21 years old or less at the child's birth (early motherhood) and 0 if she was more than 21 years old. Several studies show that the interaction and stimulation children receive from their mother are most beneficial to development if the life experience of the mother is more advanced (Corcoran, 1998; Verstraete, 2006). This variable is also measured in wave 3.

Third, marital status was transformed into a binary variable according to whether both parents were living with the child (0) or not (1) in wave 3. Substantial evidence suggests negative effects for a child's cognitive and socio-emotional outcomes under a single-parent household (McLanahan & Sandefur, 1994; Pong, 1997).

Fourth, for the total number of children in the household, we code 1 if the child has at least one sibling and 0 otherwise. This predictor was measured in wave 3. The presence or absence of other children in the household is considered because the more children in the household (regardless of the children's age), the less time and energy the parents will have to devote to a particular child. The presence of siblings can have a negative impact on the child's cognitive development (Steelman et al., 2002).

Fifth, insufficient household income was calculated as the ratio between household income and the low-income threshold in 1996 (SFR96) from Statistics Canada. Children raised in low-income families are likely to have lower math skills (Dooley & Stewart, 2004). In addition, the studies by Mayer (1997) and Blau (1999) recommend the use of permanent income as a measure (represented by the average income for all study periods). Permanent income is measured in waves 1 to 3. The individual was coded as "permanent poor" with a value of 1 if the ratio between household income and SFR96 is less than 1 for the first three waves and as 0 otherwise.

Family processes characteristics

Three parenting scales were used to measure parental behaviour. First, the positive parenting scale includes five items reflecting the frequency with which the parent compliments, plays with, laughs with, or does enjoyable activities with the child. Scores range from 0 to 20, and higher scores indicate more positive interactions. Second, the hostile/ineffective scale includes seven items in which the parent has difficulty controlling the child, disapproves of the child's behaviour, or gets angry when punishing the child. Scores range from 0 to 25, and higher scores indicate more hostile/ineffective interactions. Third, the consistent parenting scale includes five items showing the frequency with which the parent ensures that the child obeys rules or commands and the frequency with which the child gets away with behaviour for which he or she should have been punished. Scores range from 0 to 20, and higher scores indicate more consistent parenting.

Landy & Tam (1996, 1998) studied the relationship between parenting practises and cognitive, social, and behavioural outcomes and showed that positive interaction skills act as a protective factor for children at high risk. Because these variables are measured for children 2 to 11 years old, only wave 3 is used. We code 0 for children with a high score for parenting (above the 75th percentile) and 1 otherwise (except for the hostile/ineffective parenting scale, for which the opposite is true). The sub-questions used for each measure are reported in Table A.1.[3]

Table 2 shows the child, family, and parenting characteristics from the sample. We have 2,318 children who were followed over time (4 or 5 waves). All estimates are weighted. About half of the children are girls. Regarding the PPVT score, 11.70% of children had a low score (less than 85) when they were aged between 4 and 5 years. The proportion of children with behavioural problems is 31%, 26%, and 39% for hyperactivity scores, physical aggression, and

[3]Each of these scales was derived by factor analysis of parenting items in the NLSCY and has been shown to have high levels of internal consistency (Statistics Canada, 2008). See also Haeck et al. (2018) and Lebihan & Mao Takongmo (2018) for more details.

indirect aggression, respectively. Concerning family characteristics, about 10.1% of the mothers surveyed do not have a high-school diploma. Mothers who were 21 years old or less at the time of the child's birth represent 7.5% of the sample. Two-parent families represent the vast majority of the sample (about 85.8% versus 14.2% for one-parent families). About 86% of children in the study had at least one sibling. Permanently poor households represent 8.3% of the sample. The proportion of children with parenting problems is, respectively, 22%, 36%, and 18% for positive interaction, ineffective parenting style, and consistency scores.

4 Results

4.1 Identification of mathematical achievement trajectories

One of our objectives is to identify distinct mathematical abilities trajectories. In this study, the dependent variable is characterised by an individual's math scores at 7 to 15 years old. Trajectories are modelled using a censored normal distribution. Models with one to five groups are estimated. A three-group trajectory model for math achievement is selected based on BIC values: average abilities (group 1), high abilities (group 2), and low abilities (group 3) (see Figure 1). The polynomial term is quadratic for the low-abilities group and cubic for both average- and high-abilities groups.

Table 3 reports the estimates of the parameters associated with the polynomial equation between age and outcomes for each group. Trajectory 1 (i.e., the average-ability group), accounting for 47.6 % of the sample, is composed of individuals who have average mathematical skills. We observe that the math performance of these children improves as they become older (positive linear parameter of the age), but the relationship between age and math scores is not constant. In fact, the quadratic component of age is significantly negative, suggesting that the children improved less and less over time. We observe a positive change in the slope to the age of 13, suggesting that math skills improve (positive and significant cubic term). Trajectory 2 (i.e., the high-abilities group), accounting for 30.1% of the sample, is composed of individuals with particularly high mathematical skills. Again, we observe a positive - but not constant - relationship between age and math scores as a change of the slope in early adolescence (positive and significant cubic term). Trajectory 3 (i.e., the low-abilities group), accounting for 22.3 % of the sample, is composed of individuals at risk, who had lower math scores than average. Unlike the first two groups, this group is characterised by a quadratic trend.

We also note that achievement gaps among groups increase over time, especially in early adolescence. Average posterior probabilities assigned are, respectively, 0.87, 0.90, and 0.87 for groups 1, 2, and 3 and indicate a good fit of the model.

4.2 Risk factors and mathematics achievement trajectory groups

This section identifies risk factors during early childhood that affect the mathematics achievement trajectory groups of Canadian children aged 7 to 15 years. Three trajectory groups were identified: average abilities (group 1), high abilities (group 2), and low abilities (group 3).

Table 4 presents the prevalence of each risk factor by trajectory group. In general, the math performance groups differ significantly in their characteristics, except for the positive interaction score (cf, Chi-Square tests). Specifically, the groups are significantly different in terms of the sex and PPVT score of the child. We observe a greater proportion of girls and individuals with low PPVT scores in the low-abilities group than in the high-abilities group. Thus, 19.14% of children in the low-skill group had low PPVT scores when they were 4 to 5 years old; only 9.56% and 9.76% had low scores in groups 1 and 2, respectively. Similarly, the low-abilities group contains more individuals with behavioural problems (hyperactivity, physical aggression, and indirect aggression). With respect to family characteristics, the prevalence of risk factors was higher in group 3 than in the other groups. Thus, factors such as early motherhood, single-parent families, low-educated mother, and a permanently poor household are found more in the low-abilities group. One of the most significant risk factors is the mother's level of education; we observe that only 7.64% of the children in the high-abilities group have a low-educated mother, compared to 18.20% in the low-abilities group. Similarly, 8.02% of the children in the high-abilities group live with a single parent, compared to 19.11% in the low-abilities group. Group 3 also contains a higher proportion of children with a mother who was young at birth (10.67% versus 7.77% and 4.67%, respectively, for groups 1 and 2). Surprisingly, the high-ability group recorded a higher proportion of siblings (89.04% versus 83.67% and 84.94%, respectively, for groups 1 and 3). Furthermore, ineffective parenting style and lack of parental consistency seem to be determining factors for membership in the low-abilities group. In summary, we report that the trajectory groups differ mainly in PPVT score, maternal education, frequency of living in a one-parent family, and lack of consistency at home.

Table 5 reports the results of the multivariate logistic regression with the average-ability group (which contains the majority of the children) as a reference group. Log-odds ratio (estimate), standard deviations, odds ratio (OR), and 95% confidence intervals (CI) are measured. We first discuss the results for the high-abilities group, followed by the results for the low-abilities group.

Estimates show that being female reduces the likelihood of being in the high-abilities group compared to the average-abilities group (OR: 0.44; CI: [0.30; 0.63]). Having a single parent and a young mother at birth reduces the likelihood of being in the high-ability group compared to the average-ability group (OR: 0.53; CI: [0.27; 1.03] and OR: 0.49; CI: [0.23; 1.05], respectively). However, having siblings increases the likelihood of being in the high-abilities group rather than in the average group (OR: 1.68; CI: [0.31; 3.61]).

With respect to children in the low-abilities group, the two largest risk factors are a low PPVT score (OR: 2.80; CI: [1.33; 5.90]) and a low-educated mother (OR: 2.67; CI: [1.42; 4.99]). The results also indicate that problems with consistency at home increase the likelihood of belonging to the low math skills trajectory compared to the trajectory of the average-abilities group (OR: 2.44; CI: [1.41; 4.22]). However, being female decreases the odds of being in the low-skills group. We also identify some factors that do not have significant effects. For example, being in a poor family does not play a role in mathematics performance. We also find that behavioural scores have no effect on membership in the low-abilities group, nor does the number of siblings or having a young mother at birth.

5 Discussion

The objective of this study is not only to identify the mathematical abilities trajectories of Canadian children 7 to 15 years old and their probability of inclusion in each group, but also to identify the predictors during early childhood that may influence the likelihood of being in one of these abilities groups. Using group-based trajectory modelling, three trajectory groups are identified: average abilities (47.6%), high abilities (30.1%), and low abilities (22.3%). Cubic slopes are estimated for the first two groups and a quadratic slope for the last. The differences among the groups increase over time, especially in early adolescence. We then introduce risk factors during early childhood to determine which have an impact on group membership, in particular membership in the low-skill trajectory group.

Our results show that maternal education is one of the most important predictors for the low math abilities trajectory. Thus, a child with a low-educated mother is more likely to be in the low skills group than in the average group. These results are consistent with those of previous studies. PISA studies (OECD, 2004, 2007) show a positive relationship between parental education and a child's mathematics performance. Similarly, using a wide range of cognitive assessments for young children, Korenman & Winship (1995) and Currie & Thomas (1995) show that after controlling for many observable characteristics of the family and children, education and maternal skills (measured by the Armed Forces Qualification Test) are the most significant predictors in the performance of the child.

PPVT score is also a good indicator of future academic success. Indeed, children with a low PPVT score have a greater probability of being in the high-risk group than in the average group. Thus, PPVT score is a good way to detect students who might struggle in early adolescence. This confirms the results of Baker (2011) on the relationship between cognitive and behavioural development of young children and their future math performance. Similarly, Duncan et al. (2007) highlight the important contribution that reading skills and cognitive measurements at the age of 4 to 5 years make to the mathematics performance of Canadian children at the age of 8. The results obtained by Duncan et al. (2007) are also confirmed by those of Romano et al. (2010). Using NLSCY data, they show that the PPVT score is a strong predictor not only of reading skills but also of math skills when the child is 8 years old. These results reinforce the idea of "cumulative advantage processes" wherein the advantage of one person over another accumulates over time (Merton, 1973). Thus, if a child has difficulties (low academic skills) in childhood compared to others, this is likely to persist over time. The PPVT test would be a good way to detect these effects in an attempt to mitigate them.

Children at risk are also those whose parents have low parenting skills. More specifically, a lack of parental consistency leads to a higher likelihood of being in the low skills group. Several studies report that poor parenting skills are associated with lower academic achievement (Marjoribanks, 1996; Spera, 2005). Additionally, poor parenting skills lead to disorders and dysfunctions within the family dynamics, which cause disturbance for the child and consequently affect academic performance.

The sex of the child has a low impact on membership in the high-risk group. The significant effect of gender on mathematics performance is demonstrated in other studies (Caro, 2009), but we show that in the case of academic skills, gender has little impact and other aspects such as parental human capital and parental ability are most important.

Having presented the risk factors that may influence membership in the high-risk group, we can now turn our attention to other factors that are not significant. For example, being in a permanently poor family does not play a role in mathematics performance. The effects of income on the academic outcomes of children are generally weak and insignificant when compared to other factors (Blau, 1999; Dooley & Stewart, 2004). Non-monetary factors, such as maternal education, play a more important role than monetary factors in the mathematical achievements of the child.

Behavioural scores have no effect on membership in the low-abilities group. These results were also reported by Baker (2011) and Duncan et al. (2007), who found that a child's behaviour scores measured before starting school have no effect on future cognitive performance. These results are valid for any type of behaviour measured (externalised or internalised).

The number of siblings has no effect on membership in the low skills group, but it seems to have a positive effect on membership in the high skills group. This result is somewhat surprising considering that studies generally have shown that family size negatively affects a child's academic success because of the relatively smaller amount of time given to each child by the parents, as well as reduced resources allocated per child (Hanushek, 1992).

Having a mother who was young at birth has no effect on membership in the low-abilities group. This may be surprising but demonstrates the importance of the role that other factors (e.g., maternal education) play. Being in one-parent families also has no effect on membership in the low-abilities group. Thus, our study shows that risk factors such as early motherhood and single-parent families do not affect membership in the low-abilities group, but do affect whether the child will be among the best students or more in line with the average. As discussed earlier, however, maternal education is an important factor in the school performance of children.

Strengths of the study Our study has several advantages. First, by segmenting the data into multiple trajectory groups, group-based trajectory modelling provides an empirical way to summarise large amounts of data in an understandable way and introduces the risk factors that may influence membership in trajectory groups. By allowing the slopes of the trajectory groups to vary, the identification of the heterogeneity in the groups is particularly suitable (Hill et al., 2000; D. S. Nagin, 1999; D. Nagin & Tremblay, 1999). This method assumes that the population is composed of a mixture of distinct groups defined by their development trajectories. The probability of belonging to a group can then be used as the dependent variable to examine the predictors of these trajectories. The identification of the groups here is endogenous and not based on any arbitrary criterion (Côté et al., 2002). Second, the study uses a large representative sample of the Canadian population. Many early childhood predictors for the child and the child's family are measured. This creates a profile of risks that may influence the child's mathematical skills. The fact that they are measured during childhood allows for policy recommendations to be in place as soon as possible so that struggling students have the opportunity to catch up.

Limitations of the study Despite these strengths, our study also has some limitations. First, we cannot generalise the results obtained for mathematics to other cognitive tests, such as reading or vocabulary. It would also have been interesting to study whether the mathematical results obtained at ages 4 to 5 years are a more important predictor than vocabulary tests for predicting math achievements for children aged 7 to 15 years.[4] Second, other variables may also influence membership in a particular group and were not taken into account here (e.g., unobservable factors such as genetic factors, the child's motivation, and the quality of the school and the teacher). Finally, a structural analysis would be interesting to observe the mechanisms leading to these results and to improve social policies.

Implications for policy and research This study offers several policy recommendations. Indeed, the introduction of risk factors has revealed the predictors that may influence membership in specific math trajectory groups. We show that maternal education is the key consideration. Thus, we should encourage mothers to invest more in their human capital. The government should provide more funding to ensure vulnerable women have access to higher education (e.g., scholarships for student mothers and improved access to childcare).

The PPVT score is a good way to detect children at risk. Public authorities and schools should develop programs for children at risk from early childhood, such as more intensive courses in the evenings or during the holidays and more personalized teaching. This would reduce inequalities in the cognitive performance of children at the outset, before they are amplified in early adolescence.[5]

[4] Such data are lacking for each wave in the NLSCY.

[5] Increasing inequalities in cognitive performance over time is well reported by group-based trajectory modelling: The cubic terms have a positive slope for the high and middle groups later and a concave relationship for children with low abilities.

Efforts should be made not only in childhood but also in adolescence, when students begin to desire independence and to "decide" their future. Children with cognitive difficulties will remain in the background, disappointed with what society has to offer, while the more competent children will be motivated to continue on their previous trajectories.

More generally, family environment has the greatest impact on children's future. Because children spend most of their time with their families, we might consider that school compensates for some negative effects of having a mother with low education. Governments should therefore focus their efforts on improving the family environment for children at risk and increasing their exposure at school. Certainly, the government cannot significantly influence the first aspect but can influence the second; the government could increase and improve the quantity and quality of time spent in school through access to books, teaching materials, and educational activities (e.g., museum trips, sports courses).

6 Conclusion

Using the NLSCY and a group-based trajectory model, we analyse the mathematical abilities trajectories of Canadian children 7 to 15 years old. We identify three distinct mathematical abilities trajectories: average abilities (47.6%), high abilities (30.1%), and low abilities (22.3%). We also show that the gaps between trajectories increase over time, particularly in early adolescence. We analyse the links between trajectories and risk factors during childhood and find that children at risk are those with low-educated mothers, a low cognitive score at ages 4 to 5, and parents with low parenting skills.

This study allows us to develop several policy recommendations aimed at reducing the negative impact of some socio-demographic factors on the mathematical achievement of the child. This study also offers a number of future directions in continuing the study of risk factors and their effects on the academic performance of children.

References

Baker, M. (2011). Innis lecture: Universal early childhood interventions: what is the evidence base? *Canadian Journal of Economics/Revue canadienne d'économique*, *44*(4), 1069–1105.

Blau, D. M. (1999). The effect of income on child development. *Review of Economics and Statistics*, *81*(2), 261–276.

Bynner, J., & Parsons, S. (2006). Does numeracy matter more?

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian project. *Applied developmental science*, *6*(1), 42–57.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics* (Vol. 3, pp. 1801–1863). Elsevier.

Carneiro, P., Crawford, C., & Goodman, A. (2007). The impact of early cognitive and non-cognitive skills on later outcomes. CEE DP 92. *Centre for the Economics of Education.*

Caro, D. H. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education*, *32*(3), 558.

Corcoran, J. (1998). Consequences of adolescent pregnancy/parenting: A review of the literature. *Social work in health care*, *27*(2), 49–67.

Côté, S., Tremblay, R. E., Nagin, D., Zoccolillo, M., & Vitaro, F. (2002). The development of impulsivity, fearfulness, and helpfulness during childhood: Patterns of consistency and change in the trajectories of boys and girls. *Journal of Child Psychology and Psychiatry*, *43*(5), 609–618.

Côté, S., Vaillancourt, T., LeBlanc, J. C., Nagin, D. S., & Tremblay, R. E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nation wide longitudinal study of Canadian children. *Journal of abnormal child psychology*, *34*(1), 68–82.

Currie, J., & Thomas, D. (1995). *Nature vs. nurture? the bell curve and children's cognitive achievement* (Tech. Rep.). RAND-Labor and Population Program.

Dahinten, V. S., Shapka, J. D., & Willms, J. D. (2007). Adolescent children of adolescent mothers: The impact of family functioning on trajectories of development. *Journal of Youth and Adolescence*, *36*(2), 195–212.

Dooley, M., & Stewart, J. (2004). Family income and child outcomes in Canada. *Canadian Journal of Economics/Revue canadienne d'économique*, *37*(4), 898–917.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... others (2007). School readiness and later achievement. *Developmental psychology*, *43*(6), 1428.

Gagné, L. G. (2003). *Parental work, child-care use and young children's cognitive outcomes.* Statistics Canada.

Gagné, M. H., Schonert-Reichl, K., Costigan, C., Guhn, M., & Shapka, J. D. (2018). Factors predicting the stratified academic achievement trajectories of foreign-born canadian adolescents: A population-based study. *Applied Developmental Science*, 1–18.

Garon-Carrier, G., Boivin, M., Lemelin, J.-P., Kovas, Y., Parent, S., Séguin, J., … Dionne, G. (2018). Early developmental trajectories of number knowledge and math achievement from 4 to 10 years: Low-persistent profile and early-life predictors. *Journal of school psychology*, *68*, 84–98.

Green, D. A., & Riddell, W. C. (2003). Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics*, *10*(2), 165–184.

Haeck, C., Lebihan, L., & Merrigan, P. (2018). Universal child care and long-term effects on child well-being: Evidence from Canada. *Journal of Human Capital*, *12*(1), 38–98.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of political economy*, *100*(1), 84–117.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900–1902.

Herbers, J. E., Cutuli, J., Supkoff, L. M., Heistad, D., Chan, C.-K., Hinz, E., & Masten, A. S. (2012). Early reading skills and academic achievement trajectories of students facing poverty, homelessness, and high residential mobility. *Educational Researcher*, *41*(9), 366–374.

Hill, K. G., White, H. R., Chung, I.-J., Hawkins, J. D., & Catalano, R. F. (2000). Early adult outcomes of adolescent binge drinking: person-and variable-centered analyses of binge drinking trajectories. *Alcoholism: Clinical and Experimental Research*, *24*(6), 892–901.

Hoddinott, J., Lethbridge, L., & Phipps, S. (2002). *Is history destiny? resources, transitions and child education attainments in canada. final report.* ERIC.

Iruka, I. U., De Marco, A., Garrett-Peters, P., Investigators, F. L. P. K., et al. (2018). Profiles of academic/socioemotional competence: Associations with parenting, home, child care, and neighborhood. *Journal of Applied Developmental Psychology*, *54*, 1–11.

Jones, B. L., & Nagin, D. S. (2007). Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociological methods & research*, *35*(4), 542–571.

Jones, B. L., & Nagin, D. S. (2013). A note on a Stata plugin for estimating group-based trajectory models. *Sociological Methods & Research*, *42*(4), 608–613.

Jones, B. L., Nagin, D. S., & Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological methods & research*, *29*(3), 374–393.

Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child development*, *77*(1), 153–175.

Karoly, L. A., Kilburn, M. R., & Cannon, J. S. (2006). *Early childhood interventions: Proven results, future promise*. Rand Corporation.

Korenman, S., & Winship, C. (1995). *A reanalysis of the bell curve* (Tech. Rep.). National Bureau of Economic Research.

Landy, S., & Tam, K. K. (1996). Yes, parenting does make a difference to the development of children in canada. *Growing Up in Canada: National Longitudinal Survey of Children and Youth*, 117–126.

Landy, S., & Tam, K. K. (1998). *Understanding the contribution of multiple risk factors on child development at various ages*. Human Resources Development Canada, Applied Research Branch.

Lebihan, L., & Mao Takongmo, C. O. (2018). The impact of universal child benefits on family health and behaviours. *Research in Economics, Forthcoming*.

Marjoribanks, K. (1996). Family socialisation and children's school outcomes: an investigation of a parenting model. *Educational Studies*, *22*(1), 3–11.

Mayer, S. (1997). *What money can't buy*. Cambridge, MA: Harvard University Press.

McLanahan, S., & Sandefur, G. (1994). *Growing up with a single parent. what hurts, what helps.* ERIC.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Murnane, R. J., Willett, J. B., Levy, F., et al. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics*, *77*(2), 251–266.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child development*, *70*(5), 1181–1196.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological methods*, *4*(2), 139.

Nagin, D. S. (2005). *Group-based modeling of development*. Harvard University Press.

NICHD. (2002). Early child care and children's development prior to school entry: Results from the NICHD study of early child care. *American educational research journal*, *39*(1), 133–164.

OECD. (2004). *Pisa learning for tomorrow's world: First results from pisa 2003* (Vol. 659). Simon and Schuster.

OECD. (2007). *Pisa 2006: Science competencies for tomorrow's world: Volume 1: Analysis*. The PISA homepage at http:www.pisa.oecd.org.

Pong, S.-L. (1997). Family structure, school context, and eighth-grade math and reading achievement. *Journal of Marriage and the Family*, 734–746.

Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago longitudinal study. *Child development*, *69*(1), 231–246.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources*, 313–328.

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: replication and extension using a nationwide canadian survey. *Developmental Psychology*, *46*(5), 995.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics*, *86*(2), 497–513.

Shore, R. (1997). Rethinking the brain: New insights into early development.

Spera, C. (2005). A review of the relationship among parenting practices, parenting styles, and adolescent school achievement. *Educational psychology review*, *17*(2), 125–146.

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology*, *28*(1), 243–269.

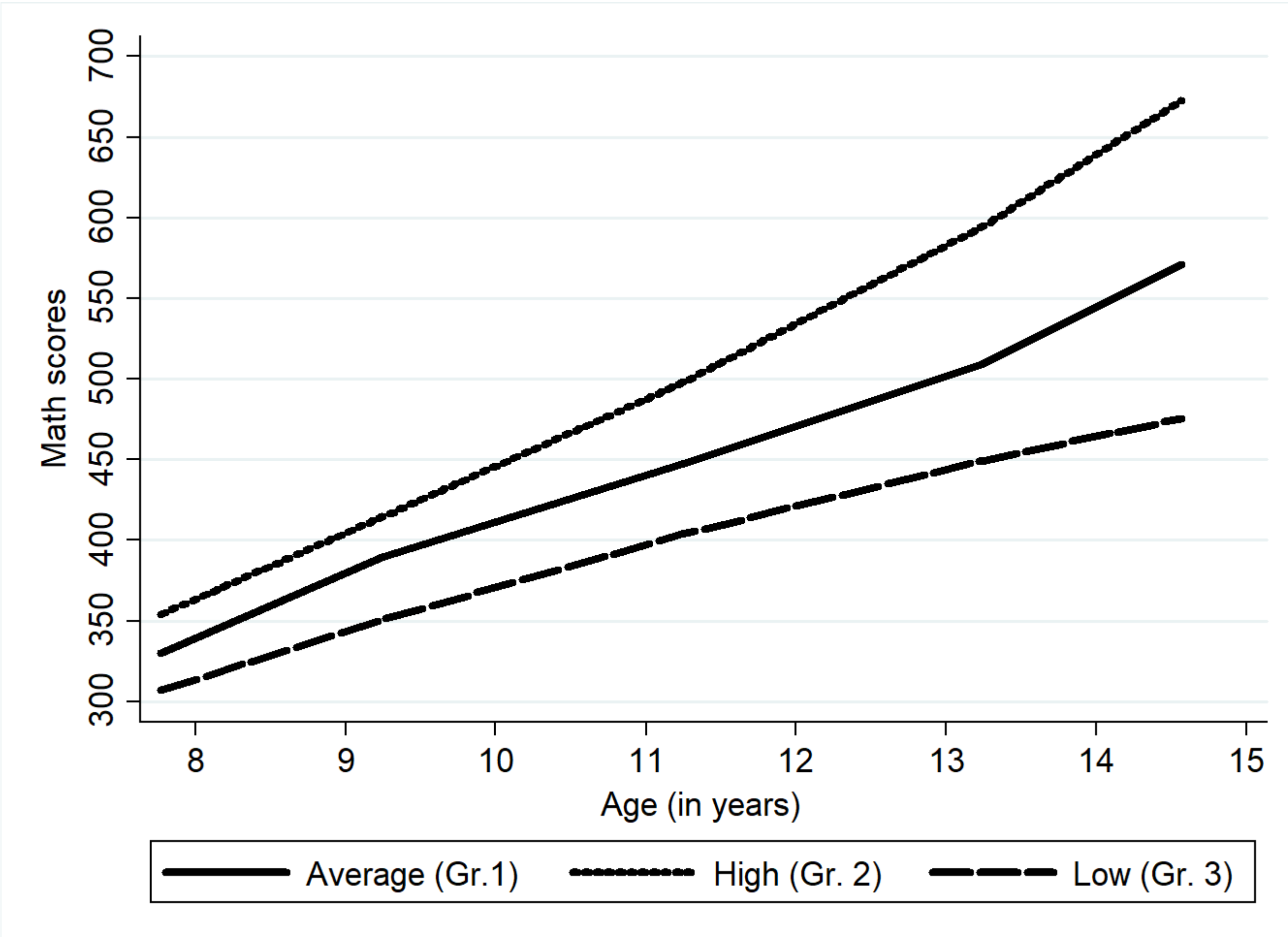
Sutcliffe, A. G., Gardiner, J., & Melhuish, E. (2017). Educational progress of looked-after children in england: a study using group trajectory analysis. *Pediatrics*, *3*, e20170503.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485).

Todd, P. E., & Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, *1*(1), 91–136.

Verstraete, M. (2006). L'impact sur le développement cognitif des enfants âgés de 4 et 5 ans de la politique de garde à 5$ du Québec.

Figure 1: Trajectories of math achievement from
7 to 15 years old



Note: Shows the trajectories of mathematics performance for children aged 7 to 15 years: average abilities (Group 1: 47.6 % of the sample); high abilities (Group 2: 30.1 % of the sample), and low abilities (Group 3: 22.3 % of the sample).

Table 1: Math score summary statistics by wave and child age

	Math Scores				
Age in years	Wave 4 (2000-01)	Wave 5 (2002-03)	Wave 6 (2004-05)	Wave 7 (2006-07)	Wave 8 (2008-09)
7	293.51				
	(38.88)				
	[709]				
8	337.81	342.35			
	(46.36)	(38.98)			
	[438]	[613]			
9	378.08	393.06			
	(49.06)	(45.28)			
	[371]	[842]			
10		418.67	412.22		
		(40.01)	(50.96)		
		[438]	[613]		
11		450.12	439.24		
		(48.18)	(51.85)		
		[371]	[857]		
12			461.42	475.3	
			(59.12)	(62.80)	
			[438]	[613]	
13			497.64	510.55	
			(63.89)	(67.78)	
			[371]	[859]	
14				549.39	587.95
				(81.21)	(88.89)
				[438]	[613]
15				579.43	603.59
				(86.12)	(83.35)
				[371]	[821]

Note: Shows the mean, standard error (in parentheses), and number of individuals (in square brackets) for standardised math scores. These statistics are classified by child age and wave and are weighted.

Table 2: Characteristics of the sample (N = 2,318 individuals)

Variables		n	%	Wave
Characteristics of child				
Sex of child	Female	1,152	49.68	3
	Male	1,166	50.32	
Low PPVT Score	Yes	271	11.70	2 and 3
	No	2,047	88.30	
Behavioural disorders				
Hyperactivity-inattention	Yes	702	31.33	3
	No	1,539	68.67	
Physical aggression	Yes	590	26.20	3
	No	1,660	73.80	
Indirect aggression	Yes	835	38.83	3
	No	1,315	61.17	
Characteristics of family				
Age of mother at birth	21 years or less	173	7.47	3
	Older than 21 years	2,145	92.53	
Family status	One parent	328	14.17	3
	Two parents	1,990	85.83	
One child	Yes	1,983	85.55	3
	No	335	14.45	
Mother graduated from high school	No	235	10.14	3
	Yes	2,083	89.86	
Permanently poor	Yes	193	8.31	1, 2 and 3
	No	2,125	91.69	
Family processes				
Positive parenting	No	515	22.22	3
	Yes	1,803	77.78	
Hostile-ineffective parenting	Yes	479	20.66	3
	No	1,839	79.34	
Consistent parenting	No	420	18.12	3
	Yes	1,898	81.88	

Notes: Shows child, family, and family process characteristics (number of individuals, proportion and wave in which the variable is measured). All risk factors are binary, and all statistics are weighted.

Table 3: Estimation of parameters

Variable	Coefficient	Standard Error
Group 1 (average abilities)		
Intercept	-823.84***	88.49
Age	292.21***	25.63
Age^2	-24.23***	2.40
Age^3	0.74***	0.07
Group 2 (high abilities)		
Intercept	-245.61**	118.37
Age	130.53***	34.03
Age^2	-9.41***	3.16
Age^3	0.33***	0.09
Group 3 (low abilities)		
Intercept	6.33	34.16
Age	46.11***	6.67
Age^2	-0.96***	0.31

Note: Shows estimated parameters and standard errors of equation between age and math scores for each group.
***: significant at 1% ; **: significant at 5%

Table 4: Sample characteristics by mathematics abilities trajectory group

Risk factors	Average ability	High ability	Low ability	Test	
	Gr.1 (n = 1,199)	Gr.2 (n = 604)	Gr.3 (n = 515)	chi2(df=2)	p
Child characteristics					
Female	54.67	43.10	47.65	8.86	0.01
Low PPVT score	9.56	9.76	19.14	34.08	0.00
Hyperactivity	32.17	27.64	34.73	19.09	0.00
Physical aggression	24.77	25.45	30.62	5.60	0.05
Indirect aggression	41.21	30.21	45.97	19.48	0.00
Family characteristics					
Low-educated mother	8.08	7.64	18.20	50.30	0.00
Early motherhood	7.77	4.67	10.67	18.32	0.00
One parent	15.78	8.02	19.11	29.26	0.00
At least one sibling	83.67	89.04	84.94	9.80	0.00
Permanently poor	8.80	5.25	11.46	21.38	0.00
Family processes					
Positive interaction	20.85	24.74	21.79	2.20	0.33
Hostile, ineffective parenting	34.71	29.51	49.04	25.01	0.00
Consistent parenting	15.40	15.57	27.76	30.57	0.00

Note: Shows the characteristics of the child, the family, and parenting by math abilities trajectory group. Chi-Square tests are performed to determine whether differences in characteristics are significant between groups. The number of degrees of freedom is 2, and p is the p-value.

Table 5: Predictors of low and high abilities trajectories.

Variables	High abilities			Low abilities		
	Estimate (SD)	OR	95% CI	Estimate (SD)	OR	95% CI
Child characteristics						
Female	-0.83***	0.44	[0.30; 0.63]	-0.46*	0.63	[0.39; 1.01]
	(0.19)			(0.24)		
Low PPVT score	0.43	154	[0.69; 3.43]	1.03***	2.80	[1.33; 5.90]
	(0.41)			(0.38)		
Hyperactivity	-0.34	0.71	[0.45; 1.12]	-0.13	0.88	[0.55; 1.41]
	(0.23)			(0.24)		
Physical aggression	0.11	1.12	[0.68; 1.82]	0.07	1.07	[0.61; 1.89]
	(0.25)			(0.29)		
Indirect aggression	-0.35	0.70	[0.44; 1.13]	0.22	1.25	[0.76; 2.03]
	(0.24)			(0.25)		
Family characteristics						
Low-educated mother	0.10	1.11	[0.54; 2.28]	0.98***	2.67	[1.42; 4.99]
	(0.37)			(0.32)		
Early motherhood	-0.72*	0.49	[0.23; 1.05]	-0.11	0.90	[0.47; 1.71]
	(0.39)			(0.33)		
One parent	-0.64*	0.53	[0.27; 1.03]	0.23	1,26	[0.61; 2.60]
	(0.34)			(0.37)		
At least one sibling	0.52*	1.68	[0.93; 3.03]	0.17	1.19	[0.60; 2.35]
	(0.30)			(0.35)		
Permanently poor	0.05	1.05	[0.31; 3.61]	0.54	1.72	[0.67; 4.40]
	(0.63)			(0.48)		
Family processes characteristics						
Hostile, ineffective parenting	-0.17	0.84	[0.55; 1.30]	0.43	1.54	[0.92; 2.56]
	(0.22)			(0.26)		
Lack of consistent parenting	0.25	1.28	[0.70; 2.36]	0.89***	2.44	[1.41; 4.22]
	(0.31)			(0.28)		
Lack of positive parenting	0.36	1.43	[0.88; 2.34]	-0.25	0.78	[0.45; 1.35]
	(0.25)			(0.28)		

Note: Shows the multinomial logistic regressions between risk factors and membership in the mathematics score trajectory groups. The reference group is the average ability group. The estimates represent the log odds ratios, and standard deviations are in parentheses. OR and CI denote the odds ratios and 95 % confidence intervals, respectively. All estimations are weighted.
***: significant at 1% ; **: significant at 5% ; *: significant at 10%

Table A.1: Risk factors index component (Appendix)

Outcome Index	**Questions**	
Hyperactivity-Inattention (Range: 0-16)	How often would you say that the child: a) Can't sit still, is restless or hyperactive?; b) Is distractible, has trouble sticking to any activity? c) Fidgets?; d) Can't concentrate, can't pay attention for long?; e) Is impulsive, acts without thinking?; f) Has difficulty awaiting turn in games or groups?; g) Cannot settle to anything for more than a few moments?; h) Is inattentive?	Almost all the time (1) to Almost never (5)
Physical Aggression (Range: 0-12)	How often would you say that the child: a) Gets into many fights? b) When another child accidentally hurts him/her, assumes that the other child meant to do it and then reacts with anger and fighting? c) Physically attacks people?; d) Threatens people?; e) Is cruel, bullies, or is mean to others?; f) Kicks, bites, hits other children?	Almost all the time (1) to Almost never (5)
Indirect Aggression (Range: 0-10)	How often would you say that the child: a) When mad at someone, tries to get others to dislike that person; b) When mad at someone, becomes friends with another as revenge? ; c) When mad at someone, says bad things behind the other's back?; d) When mad at someone, says to others: let's not be with him/her?; e) When mad at someone, tells the other one's secrets to a third person?	Almost all the time (1) to Almost never (5)
Positive interaction (Range: 0-20)	How often do you praise this child by saying something like "Good for you!" or "What a nice thing you did!" or "That's good doing!"? How often do you and this child talk or play with each other, focusing attention on each other for five minutes or more, just for fun? How often do you and this child laugh together? How often do you do something special with this child that he/she enjoys? How often do you play at sports, hobbies, or games with this child?	Never (1) to many times each day (5)
Hostile/ineffective parenting (Range: 0-25)	How often do you get annoyed with this child for saying or doing something he/she is not supposed to? Of all the times that you talk to this child about his/her behaviour, what proportion is praise?; Of all the times that you talk to this child about his/her behaviour, what proportion is disapproval? How often do you get angry when you punish this child? How often do you think that the kind of punishment you give this child depends on your mood? How often do you feel you are having problems managing this child in general? How often do you have to discipline this child repeatedly for the same thing?	Never (1) to many times each day (5)
Consistency parenting (Range: 0-20)	When you give this child a command, what proportion of the time do you make sure that he/she does it? If you tell this child he/she will get punished if he/she doesn't stop doing something and he/she keeps doing it, how often will you punish him/her? How often does this child get away with things that you feel should have been punished? How often is this child able to get out of a punishment when he/she really sets his/her mind to it? How often when you discipline this child, does he/she ignore the punishment?	Never (1) to all the time (5)