



HAL
open science

Cognitive and environmental predictors of academic skills in French middle school students

Ava Guez

► **To cite this version:**

Ava Guez. Cognitive and environmental predictors of academic skills in French middle school students. Psychology. Sorbonne Université, 2020. English. NNT : 2020SORUS240 . tel-03375054

HAL Id: tel-03375054

<https://theses.hal.science/tel-03375054>

Submitted on 12 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ecole Doctorale Cerveau Cognition Comportement (ED3C)

Laboratoire de Sciences Cognitives et Psycholinguistique (LSCP – UMR 8554)

Equipe Développement Cognitif et Pathologie

**COGNITIVE AND ENVIRONMENTAL PREDICTORS OF ACADEMIC
SKILLS IN FRENCH MIDDLE SCHOOL STUDENTS**

Ava GUEZ

Dissertation submitted for the degree of Doctor of Philosophy

Supervised by Franck RAMUS and Hugo PEYRE

Publicly defended on the 9th of October 2020

Jury:

Pascal BRESSOUX, Université Grenoble Alpes, Reviewer

Monica MELBY-LERVAG, University of Oslo, Reviewer

Timothy BATES, University of Edimburgh, Examiner

Maria MELCHIOR, Institut Pierre-Louis d'épidémiologie et de santé publique, Examiner

Lidia PANICO, Institut National d'Etudes Démographiques, Examiner

Hugo PEYRE, Université Paris Diderot, Hôpital Robert Debré, Supervisor

Franck RAMUS, CNRS, Ecole Normale Supérieure, Supervisor



Acknowledgements

Paris, le 02/10/2020

Les meilleures choses ont une fin ; certainement ce doctorat n'échappe pas à la règle. Me voici finalement au terme de ces trois ans (et dire que je ne croyais pas les gens qui me disaient que ça passerait *très* vite !) ; dans une semaine, je soutiens. Quel sentiment bizarre ! Alors que j'ai retardé le plus possible l'écriture de ces remerciements (procrastination, quand tu nous tiens !), je crois qu'il est temps de m'y mettre et d'exprimer ma gratitude envers toutes les personnes qui ont contribué à faire de ces trois années une formidable expérience professionnelle et humaine.

Je n'aurais pas pu imaginer de meilleur environnement pour réaliser une thèse que celui offert par le LSCP. Tout d'abord, je remercie chaleureusement mes deux directeurs de thèse, Franck Ramus et Hugo Peyre, qui ont si bien accompagné et aiguillé mon travail de thèse durant ces trois ans. J'ai énormément appris auprès de vous et vous en suis sincèrement reconnaissante. Franck, merci de m'avoir accueillie dans ton équipe depuis mon stage de master, merci d'avoir su partager ta rigueur scientifique, tes vastes connaissances et intérêts de recherche, ainsi que ton engagement dans la sphère publique. Hugo, merci de m'avoir initiée aux arcanes de MPlus et des SEMs (quelle fierté quand tu me poses des questions désormais !), merci pour ton soutien indéfectible et ta bonne humeur contagieuse.

Un grand merci également à tous les membres de l'équipe Pathologie et développement cognitifs. Merci aux super doctorantes, ingénieure de recherche, chercheuse émérite et post-docs de l'équipe, passés et présents : Alice, Camille, Ghislaine, Marie-Christine, Mikel et Raphaël, pour les réunions d'équipe et discussions stimulantes, vos sourires et vos conseils avisés. Merci particulièrement à Alice, ma grande sœur de thèse, pour toutes nos pauses thé, ton écoute et soutien précieux ! Merci à mes stagiaires pour leur fraîcheur et leur grande aide : Patricia et Sami, qui ont donné de leur temps pour écouter et coder les enregistrements des enfants EDEN ; Baptiste, Cécile, Elisa et Valentin, qui m'ont permis d'explorer de nouveaux sujets de recherche parallèlement à mes travaux de thèse.

Qu'auraient été mes trois ans sans notre super salle #salleC402 ? Merci à Alice, Camila, Camille, Georgia, Ghislaine, Leticia, Mélissa, Monica et Si d'avoir rendu cette salle si vivante et accueillante. Merci pour vos sourires, votre bonne humeur, les repas, goûters, verres et rires partagés, vos conseils et soutien ! Merci à tous les autres collègues du labo qui ont contribué à rendre ces trois années aussi sympathiques et enrichissantes : Cécile, Charlotte, Gerda, Hualin, Matthieu, Mireille, Naomi, et j'en oublie... Merci à Radhia, Vireack et Michel d'avoir été là pour toute question administrative ou technique et pour votre présence rassurante. Et bien sûr, merci à nos deux directrices successives du labo, Anne Christophe et Alex Cristia, qui ont également toutes les deux été mes tutrices de master et doctorat. Merci pour votre bienveillance, et merci d'avoir fait du labo ce lieu d'échange convivial, cette grande famille. Alex, merci également de m'avoir fait confiance et permis de collaborer à ce passionnant projet qu'est l'analyse des données LENA à travers le monde.

Mes remerciements s'étendent également en dehors du labo ! I sincerely thank Timothy Bates, Pascal Bressoux, Monica Melby-Lervåg, Maria Melchior, and Lidia Panico for doing me the honor of accepting to read, review and discuss my dissertation. I am particularly grateful to Pascal Bressoux and Maria Melchior, who were also part of my PhD committee and followed my work since the first year. Merci pour vos précieux conseils.

Je remercie également les différents co-auteurs qui ont contribué à cette thèse : Cécile Di Folco, Nicolas Gauvrit, Barbara Heude, Marion Le Cam, Manuela Piazza et Pedro Pinheiro-Chagas. Merci tout particulièrement à Manuela de m'avoir si bien accueillie ces quelques jours à Rovereto et d'avoir partagé avec moi ses connaissances en cognition mathématique.

Sur une note plus personnelle, je tiens à exprimer un énorme merci à ma famille, soutien des premières heures. Merci pour votre amour, vos conseils et vos encouragements ; je ne serai pas là sans vous. Et Gaurav: thank you for believing in me more than myself, for pushing me and challenging me, for being such an immense source of comfort when needed, and for being the best ppt adviser!

Pour finir, je remercie l'ED3C et Sorbonne Université d'avoir financé cette thèse !

Abstract

Learning to read, write and count is fundamental to unlock one's potential and flourish in modern society. These essential skills are acquired throughout schooling, building on a rich set of cognitive and socio-emotional skills, and influenced by a myriad of environmental and genetic factors. This dissertation aimed at providing a greater understanding of what fosters or hampers the acquisition of academic skills in children. It has two leading threads: first, the necessity to portray and assess the multidimensional factors influencing academic outcomes: cognitive, socio-emotional, and environmental; second, the need to assess these in understudied contexts, here France. To do so, we have conducted a series of studies using longitudinal data from two French cohorts, the EDEN cohort and the DEPP Panel 2007, assessing the relative influences of a wide variety of factors on diverse aspects of academic achievement in middle school.

First, we studied the extent of the association between intelligence and academic skills in France. We assessed the strength of this relationship, as well as the socio-economic and conative influences on academic skills and their progression beyond the role of IQ. We further investigated the relationship between IQ and academic achievement among intellectually gifted student. Second, we dug into one component of academic skills, numeracy, examining its preschool predictors. We assessed the relative predictive power of cognitive, socio-emotional and environmental factors on arithmetic skills as well as their mediation relationships; and investigated the differential cognitive predictors of addition, subtraction and multiplication. Third, we similarly studied the preschool cognitive, socio-emotional and environmental influences on the acquisition of different literacy skills, and their mediating relationships. Fourth and last, we examined sex differences in both literacy and numeracy, assessing the influence of evaluation characteristics on these gaps.

The results of these studies provide valuable insights into the mechanisms underlying the acquisition of academic skills in France, and have practical implications for practitioners and actors in the education sphere.

Résumé

Apprendre à lire, écrire et compter est indispensable pour exprimer son potentiel et s'épanouir dans la société actuelle. Ces compétences essentielles, acquises tout au long de la scolarité, se fondent sur un socle riche de compétences cognitives et socio-émotionnelles, et sont influencées par une myriade de facteurs environnementaux et génétiques. Cette thèse vise à fournir une meilleure compréhension des conditions et attributs qui favorisent ou entravent l'acquisition des apprentissages scolaires chez les enfants. Elle s'articule autour de deux axes principaux : premièrement, la nécessité de décrire et d'évaluer les facteurs multidimensionnels qui influencent les résultats scolaires – cognitifs, socio-émotionnels et environnementaux ; deuxièmement, le besoin d'estimer le rôle de ces facteurs dans des contextes peu étudiés, ici la France. Pour ce faire, nous avons mené une série d'études s'appuyant sur les données longitudinales de deux cohortes françaises, la cohorte EDEN et le Panel 2007 de la DEPP, évaluant les influences relatives d'une large gamme de facteurs sur divers aspects de la réussite scolaire au collège.

En premier lieu, nous avons étudié la relation entre intelligence et compétences scolaires en France. Nous avons évalué la magnitude de cette association, ainsi que les prédicteurs socio-économiques et cognitifs de la réussite scolaire et de sa progression au-delà du rôle du QI. Nous avons ensuite étudié cette relation chez les élèves à haut potentiel intellectuel. En deuxième lieu, nous nous sommes penchés sur l'une des composantes des apprentissages scolaires, le calcul. Nous avons évalué le pouvoir prédictif relatif des facteurs cognitifs, socio-émotionnels et environnementaux précoces sur les compétences en arithmétique à 11 ans et demi, ainsi que leurs relations de médiation ; puis nous avons examiné les prédicteurs cognitifs différentiels de l'addition, de la soustraction et de la multiplication. En troisième lieu, nous avons étudié de manière similaire les influences des facteurs cognitifs, socio-émotionnels et environnementaux précoces sur l'acquisition de la lecture et de l'orthographe, ainsi que leurs relations de médiation. Pour finir, nous avons analysé les différences entre filles et garçons en Français et mathématiques, en estimant l'influence du type d'évaluation sur ces écarts.

Les résultats de ces études élargissent les connaissances des mécanismes qui sous-tendent l'acquisition des apprentissages scolaires en France, et ont des implications pratiques pour les professionnels et acteurs du domaine de l'éducation.

Table of contents

Acknowledgements	3
Abstract	5
Résumé	6
Table of contents	7
Publications	10
Chapter 1 - General introduction	13
1.1 What shapes early cognitive development	14
1.1.1 Main predictors of cognitive development	14
1.1.2 The importance of controlling for confounding variables	17
1.1.3 From simple associations to complex relationships	20
1.2 Acquiring academic skills: building blocks and determinants	24
1.2.1 The role of cognitive abilities	25
1.2.2 The role of social, behavioral and emotional skills	29
1.2.3 The role of socio-demographic factors	30
1.3 Studying the acquisition of academic skills in France	32
1.3.1 Learning in France: an outlook	33
1.3.2 Data: two French cohort studies	35
1.3.3 Objectives and research questions of this dissertation	38
Chapter 2 - Intelligence and academic skills	41
2.1 Predictors of the IQ-achievement gap	42
2.1.1 Abstract	42
2.1.2 Introduction	43
2.1.3 Method	46
2.1.4 Results	54
2.1.5 Discussion	60
2.1.6 Appendix A: Data	66
2.1.7 Appendix B : Results	69
2.2 Are high-IQ students more at risk of school failure?	73
2.2.1 Abstract	73
2.2.2 Introduction	74
2.2.3 Method	77
2.2.4 Results	83
2.2.5 Discussion	93

Chapter 3 - Early predictors of arithmetic skills	97
3.1 Cognitive and environmental predictors of problem solving skills	98
3.1.1 Abstract	98
3.1.2 Introduction	99
3.1.3 Method	102
3.1.4 Results	110
3.1.5 Discussion	119
3.1.6 Appendix	124
3.2 Cognitive predictors of multiplication, addition and subtraction	127
3.2.1 Abstract	127
3.2.2 Introduction	128
3.2.3 Method	130
3.2.4 Results	134
3.2.5 Discussion	138
Chapter 4 - Early predictors of literacy skills	141
4.1 Cognitive and environmental predictors of reading and spelling	141
4.1.1 Abstract	141
4.1.2 Introduction	142
4.1.3 Method	149
4.1.4 Results	157
4.1.5 Discussion	172
4.1.6 Appendix	179
Chapter 5 - Sex differences in academic skills	182
5.1 Sex differences are modulated by evaluation type	182
5.1.1 Abstract	182
5.1.2 Introduction	183
5.1.3 Method	185
5.1.4 Results	191
5.1.5 Discussion	196
5.1.6 Appendix	200
Chapter 6 - General discussion	204
6.1 Cognitive and socio-emotional foundations of academic skills	205
6.1.1 Cognitive skills	205
6.1.2 Socio-emotional skills	207
6.2 Environmental and individual influences on academic achievement	208
6.2.1 Parental socio-economic and cultural factors	208
6.2.2 Sex	210
6.2.3 Pre-natal and birth factors	211
6.3 General limitations	211

6.3.1	Genetic confounding	212
6.3.2	Correlation and causality	213
6.4	Practical implications and conclusion	214
References		216

Publications

Articles included in the dissertation

Guez, A., Panaïotis, T., Peyre, H., & Ramus, F. (2018). Predictors of the IQ-achievement gap in France: A longitudinal analysis. *Intelligence*, 69, 104–116.

Guez, A., Peyre, H., Le Cam, M., Gauvrit, N., & Ramus, F. (2018). Are high-IQ students more at risk of school failure? *Intelligence*, 71, 32–40.

Guez, A., Peyre, H., & Ramus, F. (*under review*). Sex differences in academic achievement are modulated by test type.

Guez, A., Peyre, H., Pinheiro-Chagas, P., Heude, B., Piazza, M., & Ramus, F. (*under review*). Early cognitive, emotional and environmental predictors of arithmetic abilities in 11.5-year-old children.

Guez, A., Pinheiro-Chagas, P., Peyre, H., Ramus, F. & Piazza, M. (*in prep.*). Differential cognitive predictors of multiplication, addition and subtraction skills in 11.5-year-old children.

Guez, A.*, Di Folco, C.*, Peyre, H., & Ramus, F. (*in prep.*). Early cognitive and environmental predictors of literacy skills at 11.5 years old.

Guez, A., Peyre, H., Williams, C. M., Labouret, G., & Ramus, F. (*in prep.*). The epidemiology of cognitive development

Other articles and contributions

Guez, A. & Ramus, F. (2019) Les écrans ont-ils un effet causal sur le développement cognitif des enfants ? *Revue suisse de pédagogie spécialisée*, 4, 14-21.

Gauvrit, N. & **Guez, A.** (2018). Réussite scolaire et professionnelle des personnes à haut potentiel intellectuel. *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 154, 313-318.

Guez, A., Peyre, H. & Ramus, F. (*in prep.*). Bidirectional relationships between language skills and behavioral difficulties during childhood.

Di Folco, C.*, **Guez, A.***, Peyre, H. & Ramus, F. (*in prep.*). Prevalence of developmental dyslexia: A comparison of DSM-5 and ICD-11 criteria.

Moreau-Pernet, B., **Guez, A.**, & Ramus, F. (*in prep.*). Can grade repetition benefit academic achievement?

Bernu, V., **Guez, A.**, Peyre, H., & Ramus, F. (*in prep.*). Are sex differences in cognitive development moderated by SES?

Bergelson, E.,...,**Guez, A.**,..., & Cristia, A. (*in prep.*). Everyday language input and learning in 1,000 children from 6 continents.

Tang, J., Su, Y. Yao, Y., Peyre, H., **Guez, A.**, Zhao, J. (*under review*). Why do Girls Perform Worse than Boys in Math? The Affective Effect of Mathematical Anxiety.

* The authors contributed equally to this work

Chapter 1 - General introduction

Mastering basic academic skills (literacy and numeracy) is crucial to be wholly involved in our modern society and live up to one's full potential. Notably, reading, writing and counting skills are associated with important life outcomes such as higher educational qualifications (Duncan et al., 1994), higher socio-economic status (SES) in adult life (Ritchie & Bates, 2013), and better health (Ancker & Kaufman, 2007; Berkman et al., 2011). In OECD countries, more than 20% of 15-year-olds still have not attained the sufficient proficiency in reading to participate fully in society, and the same goes for mathematics (OECD, 2016a, 2019). In France, very similar proportions are observed (OECD, 2016a, 2019). Designing effective teaching programs and pedagogical interventions to tackle this situation requires a deep understanding of the set of factors which contribute to individual differences in the acquisition of these basic academic skills. Learning to read, write and count relies heavily on a broad set of cognitive abilities that develop during fetal maturation and throughout childhood. As such, they can be seen as an integral part of human cognitive development, their mastery being a particular milestone in a “developmental continuum”¹. Understanding what makes a child a better learner thus entails identifying and disentangling the intricate web of pre- and post-natal factors which contribute to cognitive development early on in childhood, in which ways early cognitive development lays the foundations for the acquisition of basic academic skills, and how environmental factors affect academic outcomes beyond their influence on cognitive development. This dissertation falls within this endeavor, focusing specifically on the French context. In this introductory section, we first expose the known pre- and post-natal predictors of cognitive development and their complex relationships (Part 1.1). We then describe how specific aspects of cognitive development ground and foster the acquisition of academic skills, as well as the additional influence of socio-demographic factors (Part 1.2). Lastly, we expound the context and methodology of the present research work, i.e. the study of the cognitive and environmental predictors of academic skills in France based on the longitudinal analysis of two cohort studies (Part 1.3).

¹ We borrow this terminology from Whitehurst & Lonigan (1998), who developed the idea that “the acquisition of literacy is best conceptualized as a developmental continuum, with its origins early in the life of a child, rather than an all-or-none phenomenon that begins when children start school”.

1.1 What shapes early cognitive development²

Cognitive development is the result of an intricate combination of genetic and environmental factors. Decades of research in psychology and epidemiology have highlighted the extent to which these factors predict variance in cognitive abilities and traits (1.1.1), and made considerable advances in uncovering the complexity of their relationships, from disentangling their unique contributions (1.1.2) to mapping out their convoluted web of interactions (1.1.3).

1.1.1 Main predictors of cognitive development

Cognition is a vast array of abilities and traits, including a variety of domains – such as language, motor skills, reasoning, working memory or attention, but also social, emotional and behavioral skills. Providing a detailed picture of the multitude of factors that influence these various areas is no easy task. We do not intend here to be exhaustive, but rather to report the main factors which have stood out in the literature, and for which solid evidence – from meta-analyses when possible, and large cohort studies otherwise – has been provided. In doing so, we chose not to restrict our scope to some specific domains of cognition in order to illustrate the diversity of influences that predictors can have on different areas of cognitive development.

Sex

Sex differences in cognitive development have been the focus of a wide range of studies in psychology. Although male and female children are largely similar, they show some differences in the developmental trajectories of certain cognitive functions. For instance, while there is no sex difference in general intelligence (Deary et al., 2007), robust small differences have been found in specific cognitive abilities. Hence, meta-analytic evidence shows that girls have better verbal skills than boys (Hyde & Linn, 1988), and boys perform better in mental rotation tasks (Maeda & Yoon, 2013; Voyer et al., 1995). Besides, boys tend to be more at risks of having neurodevelopmental disorders (May et al., 2019)

² This section is based on part of the following article: **Guez, A., Peyre, H., Williams, C. M., Labouret, G., & Ramus, F. (in prep.).** The epidemiology of cognitive development.

such as ASD (male-to-female ratio equal to 3:1, Loomes et al., 2017) and ADHD (3:1, Willcutt, 2012).

Prenatal exposure

Exposure to certain elements during pregnancy may have detrimental effects on the child's cognitive development. For instance, children of epileptic mothers who have been exposed in utero to valproic acid, an antiepileptic drug, have on average lower general cognitive abilities (Banach et al., 2010) and are at higher risks of developing ASD (Christensen et al., 2013) and various neurodevelopmental problems (Blotière et al., 2020). Prenatal alcohol exposure also has negative consequences for cognitive development: meta-analytic studies suggest that moderate alcohol intake during pregnancy (3-6 drinks per week) is negatively associated with child behavior outcomes ($d=-0.15$; Flak et al., 2014), while binge drinking (more than 4 drinks per occasion) and heavy drinking (more than 2 drinks per day) are negatively associated with general cognitive development (respectively $d=-0.13$ for binge drinking, Flak et al., 2014; and $d=-0.53$ for heavy drinking, Testa, 2003).

Birth factors

Cognitive development is also associated with several birth characteristics, such as gestational age, birth weight and the Apgar score (which evaluates a newborn's clinical status). Thus, preterm children and those with low birth weights experience a variety of cognitive deficiencies including linguistic, intelligence, sensory, and motor difficulties, compared to term children with normal birth weights (Aarnoudse-Moens et al., 2009; Barre et al., 2011; Beauregard et al., 2018; Courchia et al., 2019; de Kieviet et al., 2009; Nguyen et al., 2018; Twilhaar et al., 2018). Cognitive declines range from 0.2 to 0.3 SD for preterm (<37 weeks) and early term children (37–38 weeks; Beauregard et al., 2018) and up to 0.86 SD for very preterm children (<32 weeks) and/or with very low birth weight (<1500 g) (Twilhaar et al., 2018). Lastly, a low Apgar score (<7) is additionally associated with lower cognitive abilities (Ehrenstein, 2009; Razaz et al., 2016) and teacher-rated hyperactivity and inattention (Guhn et al., 2020; Razaz et al., 2016).

Parental and social factors

Parental and social factors after birth also explain individual differences in children's cognitive abilities and traits. For instance, breastfeeding is associated with higher offspring's general cognitive abilities (Horta et al., 2015: 3.44 more IQ points) and with lower risks of developing certain behavioral problems such as ADHD symptoms (Tseng et al., 2019: Odds ratio for non-breastfeeding = 3.71). Other parental characteristics such as paternal and maternal age, socio-economic status, and maternal depression, have also been found to be associated with children's cognitive outcomes in large scale cohort studies. Younger maternal age is negatively linked with general cognitive abilities (less than 25 years old: $\beta = -0.13$ to -0.17), while the reverse is true for older age ranges (35-39 years old: $\beta = 0.10$) (Goisis et al., 2017). However, advanced maternal age is positively associated with an increased likelihood of internalizing problems (OR=1.06), but negatively linked with externalizing problems (OR=0.88) (Saha et al., 2009). On the contrary, paternal age is likely to increase the risks of developing externalizing problems (OR=1.12) but not internalizing problems (Saha et al., 2009). Both advanced maternal and paternal age (older than 40) are associated with augmented risk of developing ASD (Reichenberg et al., 2006; Sandin et al., 2017). Besides, parents' socio-economic status is positively associated with a wide range of cognitive outcomes – general cognitive ability, executive functions, behavioral outcomes, language development –, as supported by meta-analytic evidence (Lawson et al., 2018; Letourneau et al., 2013; Scaff & Cristia, in prep.). Lastly, education is an obvious contributor to children's cognitive abilities, with an increase of, on average, 3.4 IQ points for 1 year of education (Ritchie & Tucker-Drob, 2018).

Genes

A survey of the factors influencing cognitive development would not be complete without a mention of genetic factors. The collective effect of genes on cognition has been investigated in heritability studies, which determine the share of variance in phenotypes that is due to genetic variance. Historically, such studies have relied on the comparison between mono- and di-zygotic twins (Bartels et al., 2002; Bishop et al., 1995), but have also exploited other situations such as adoption at birth and more generally trait correlations between relatives of varying genetic and environmental similarity (Plomin et al., 1997). Across all cognitive traits, heritability has typically been found to lie between 20 and 80% (Plomin et al., 1994), making the genome the single most important factor in cognitive

development (although each individual genetic variant only has a minute effect on cognition).

Since the beginning of the 21st century, new molecular genetic methods have complemented twin and family studies. Genome-wide complex trait association (GCTA) studies use whole-genome analysis to estimate the proportion of phenotypic variance that can be explained by genetic variance, directly measured across dozens or hundreds of thousands of single-nucleotide polymorphisms (SNPs) (Yang et al., 2011). For reasons that are well understood, they show systematically lower heritability estimates than twin studies (Trzaskowski et al., 2013), but do confirm the substantial influence of genes on most cognitive traits. Such genome-wide association results are now being used to compute polygenic scores, which cumulate the predictive power of thousands of SNPs that are most strongly associated with the phenotype of interest. Current polygenic scores may account for up to 10% of the variance in cognitive performance (Lee et al., 2018).

Beyond documenting the contribution of genes to cognitive development, perhaps the most interesting contribution of such genetic studies is to enrich studies of environmental factors by allowing one to consider interactions between genetic and environmental factors, and providing a way to adjust for the confounding effects of genetic factors on environmental ones.

1.1.2 The importance of controlling for confounding variables

As one can imagine, many of the various predictors of cognitive development are correlated with each other. It is therefore often necessary to measure as many factors as possible, and adjust them on one another to identify the specific contribution of each one (see **Figure 1**). For instance, parental education is correlated with family income, with the quality of medical care, and with parent/child interactions. Failing to measure and control for any of these factors may lead to overstating the influence, or misattributing a causal role, to the others. We develop and illustrate two main types of such confounding: confounding due to omitted environmental variables, and to genetic factors (i.e. gene-environment correlations).

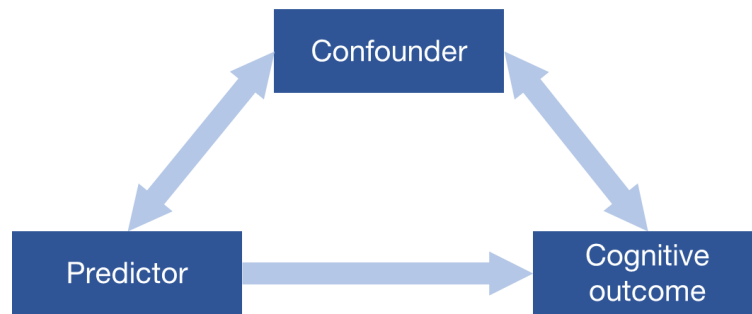


Figure 1: Confounding variables. Omitting to account for a confounder (a variable correlated with both the predictor of interest and the cognitive outcome) results in a biased estimate of the association between this predictor and the outcome.

Environmental confounders

Environmental factors are the vast array of conditions to which a child can be exposed during or after pregnancy. It is often the case that environmental predictors of cognitive development are correlated with one another. In order to identify their unique contributions and compare their relative influence, it is therefore important to control for the influences of potential confounders. One example of such confounded relationship is that between breastfeeding and maternal IQ. Breastfeeding has been purported to have a positive influence on cognitive development due to the particular composition of maternal milk. However, when controlling for maternal IQ, the association between breastfeeding and the child's intelligence falls from 3.44 IQ points to 2.62 (Horta et al., 2015). Similarly, after matching breastfed children with non-breastfed children on a range of individual and parental characteristics, the difference in general cognitive outcomes considerably shrinks and becomes non-significant (Girard et al., 2018). These results suggest that a large part of the association between breastfeeding and the child's cognitive development may stem from richer mother-child interactions, rather than nutritional benefits. In a similar fashion, maternal smoking during pregnancy has long been believed to be associated with decreased cognitive outcomes. However, large scale studies which controlled for a wide range of factors thought to be correlated both with maternal smoking and cognitive outcomes, such as maternal education, found no evidence for such association (Batty et al., 2007; Gilman

et al., 2008). Therefore, maternal smoking in itself does not seem to be detrimental to the infant's cognitive development.

Gene-environment correlations

While it is commonplace in social science and epidemiological research to measure and control as many potentially confounding factors as possible, this approach is often restricted to environmental factors. Yet, genetic factors are also often intertwined with environmental factors. This has been known for a long time, with the paradoxical discovery of the heritability of environmental factors, also known as “the nature of nurture” (Plomin & Bergeman, 1991). For instance, the very exposure to life events (accidents and trauma), an unambiguous environmental factor, is more concordant between monozygotic than between dizygotic twins, hence has a non-null heritability. This can be understood as reflecting genetic influences on cognitive traits such as risk-taking or impulse control, or less directly, genetic influences on intelligence which in turn has an effect on the likelihood of understanding and following basic safety recommendations. Gene-environment correlations can take different forms (Pingault et al., 2018; Rutter, 2007). They can be passive, such as when parents with language skills both genetically transmit these predispositions to their children, and provide a richer linguistic environment for these children to grow up in. They can be evocative, such as when children with good language learning predispositions talk more and better, and therefore elicit richer language input in return. They can also be active, such as when children with good language learning predispositions actively seek peers with good verbal skills, books, and challenging linguistic environments. In all cases, studies may measure the association between the linguistic environment and children's language abilities, and make incorrect (or inflated) causal inferences if they don't control for genetic transmission.

Nowadays, molecular genetics offers a way to directly measure and control genetic influences. It has indeed been shown that certain polygenic scores are significantly correlated with some environmental factors known to have an effect on cognitive development. For instance, a child's genome-wide polygenic score (GPS) for educational attainment is correlated with parental education, income, and age at the child's birth, with number of books in the home, with breastfeeding duration, with smoking during pregnancy, with whether the TV is usually on, with smacking or slapping (Krapohl et al., 2017). Although parental SES is one of the main predictors of educational achievement, this

relationship may be to a large extent accounted for by genetic variance (Trzaskowski et al., 2014).

1.1.3 From simple associations to complex relationships

More often than not, it is plausible to imagine that predictors and cognitive outcomes may be related in more complex ways than portrayed so far. What is more, cognitive skills are also related to each other due to the dynamic nature of cognitive development – it is the idea that “skills beget skills” (Cunha & Heckman, 2007). More sophisticated statistical models can be used to understand the particular mechanisms through which a predictor ultimately affects cognitive outcomes: is the effect mediated through a third factor? Does the effect depend on particular circumstances? To what extent do two factors exert reciprocal influences on each other?

Mediation effects

In order to have a more complete picture of the effect of one factor on cognitive development, one can look at the potential mediators of such relationship. Mediation effects designate a relationship (thought as causal) between one distal factor, such as parental education, one proximal factor, such as breastfeeding, and one outcome, such as verbal cognitive ability (see **Figure 2**): the positive association between parental education and a child’s verbal skills is partly explained by the fact that higher educated mothers breastfeed more, which is itself associated with higher verbal outcomes (Peyre, et al., 2016). Statistical models allow to estimate to what extent the effect of the distal factor is mediated through the proximal factor. Different methods can be used, depending on the nature of the variables. When the relationships between variables is linear, that the variables are normally distributed and that there are no interactions, structural equation models (SEM) are one efficient way to estimate mediation effects, even when multiple mediators are present (Preacher & Hayes, 2008). When such assumptions are not reasonable, SEMs can be used in an exploratory fashion to generate hypotheses, but will often need to be followed by more rigorous analysis strategies. Causal mediation analysis, a method based on counterfactual reasoning, provides a more rigorous framework for estimating such relationships (VanderWeele, 2016).

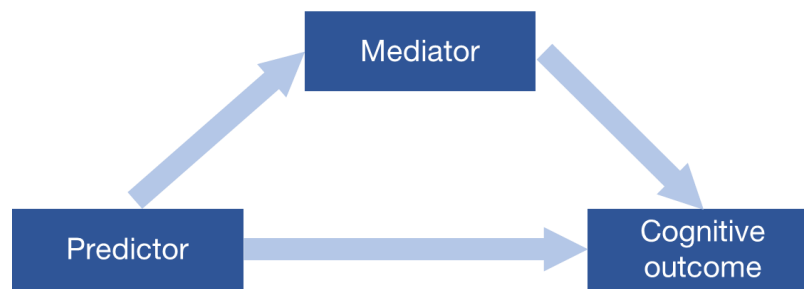


Figure 2: Mediation relationships. The underlying mechanism through which a predictor ultimately influences cognitive outcomes may be captured by a mediation relationship, whereby the predictor’s influence is partly (or fully) explained by its effect on a third, mediator variable which is also associated with cognitive outcomes.

Moderation/interaction effects

Given exposure to similar environmental factors, distinct individuals may react differently. This may be due to different developmental history, sex, or genes conferring different vulnerability or potential. This phenomenon is known as moderation of the effect of one factor by another, or interaction effect between the two predictors: i.e., when the effect of one factor depends on the presence or the value of another (see **Figure 3**). For instance, the negative influence of prenatal alcohol exposure on the child’s executive functions is greater when the mother is older (Burden et al., 2005; Chiodo et al., 2010).

Beyond environmental factors interacting with each other, the child’s sex seems to moderate the effects of certain environmental factors on cognitive development. For instance, low birth weight is a long-term risk factor for depression in adolescent girls, but not in boys, and only in conjunction with other childhood risk factors (Costello et al., 2007). As another example, the well-known male advantage in spatial skills has been found to emerge only at middle/high SES, but not at low SES, thus constraining the possible explanations for this sex difference (Levine et al., 2005).

Genetic makeup has also been shown to interact with environmental factors. Understanding such interactions may throw new light on well-established environmental effects. For instance, it has long been known that childhood maltreatment is associated with conduct disorder and with later antisocial personality behavior. This may be interpreted as reflecting a form of learning by imitation. However, not all maltreated children become

maltreating parents. In a landmark study, Caspi et al. (2002) showed that a particular polymorphism of monoamine oxidase a (MAO-A) interacted with childhood maltreatment, such that carriers of the low protein expression variant were more at risk of developing conduct disorder if they were maltreated (but not if they were not). This result, strengthened by meta-analytic evidence (Byrd & Manuck, 2014), suggests that the learning-by-imitation interpretation is at best incomplete. Thus, given the variations in the response to environmental factors, it is important to consider genetic factors as one possible source of this variability.

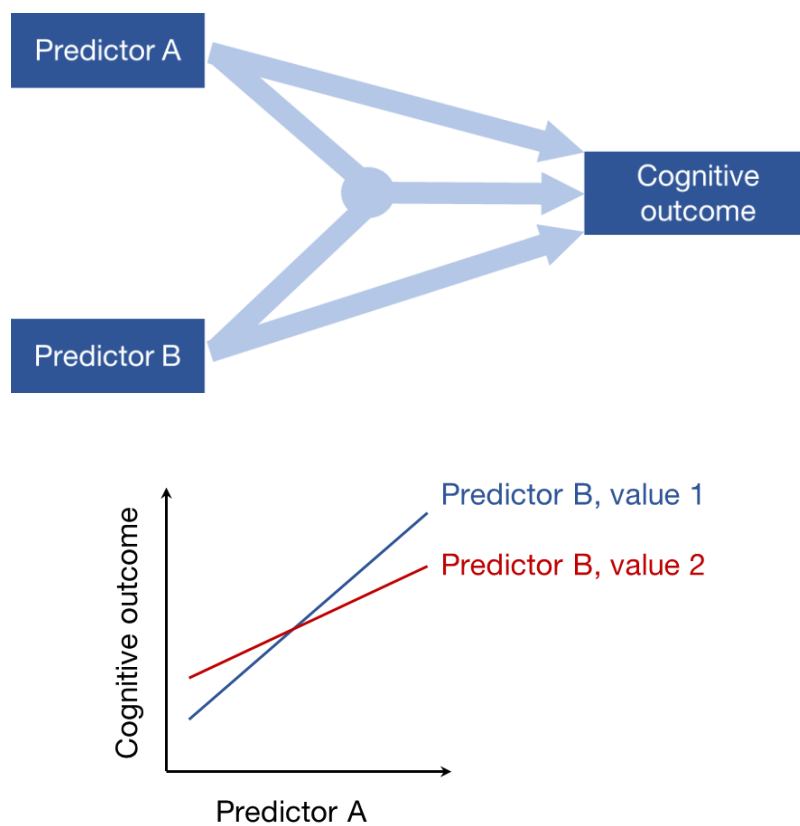


Figure 3: Moderation/interaction effects. The effects of two predictors on cognitive outcomes can be multiplicative, when the effect of a predictor A varies with the values of a second predictor B: the influence of A is moderated by B. The bottom panel illustrates a moderation/interaction relationship where the positive relationship between a continuous predictor A and the outcome is more or less strong depending on the value of a categorical predictor B.

Reciprocal relationships

When two variables which evolve in time are correlated, it is often hard to determine which one causes the other, or, if each has a causal effect on the other, which one has the larger effect (see **Figure 4**). For example, language abilities and behavioral problems are two cognitive outcomes which are correlated, but for which the direction of the relationship is not obvious: it is possible that early behavioral problems impair language development, but also that early language difficulties prevent children from properly regulating their behavior. It is also possible that there is no causal link between the two outcomes, but that both are caused by a third, potentially unobserved, factor, which creates a correlation between them. Cross-lagged panel models are a kind of structural equation models which can help disentangling such longitudinal relationships. In these models, the two variables are measured at different points in time and are simultaneously regressed on past values of themselves and on past values of the other one. When measures are available at more than two time points, more sophisticated models can be used, which allow distinguishing between-person from within-person variance; for example, models including a random intercept (Hamaker et al., 2015). These models are able to estimate to what extent a variable A affects the within-person change in variable B, and vice-versa. Cross-lagged panel models examining the relationships between language abilities and ADHD symptoms have thus shown that better early language skills prevent the development of ADHD symptoms, but that early ADHD symptoms do not impair language acquisition (Petersen et al., 2013; Peyre et al., 2016).

Similar methods can be applied when exposure to a risk factor varies with time, and its relationships with cognitive outcomes are thus unclear. For example, exposure to screens is correlated with children's cognitive abilities (Madigan et al., 2020; Walsh et al., 2018). However, we do not know a priori if this correlation conceals a causal relationship from screen time to cognitive abilities (for example, if watching TV or playing a video game deters children from doing activities more beneficial to cognitive development), a causal relationship from cognitive abilities to screen time (for example, if children with lower cognitive abilities are more attracted to screens), or is simply due to external factors (for example, if children from lower socioeconomic backgrounds both have lower cognitive abilities and are more exposed to screens). Going beyond simple associations by using a random intercept cross-lagged panel models showed that there is a small negative link from screen time to general cognitive development, but not the reverse (Madigan et al., 2019).

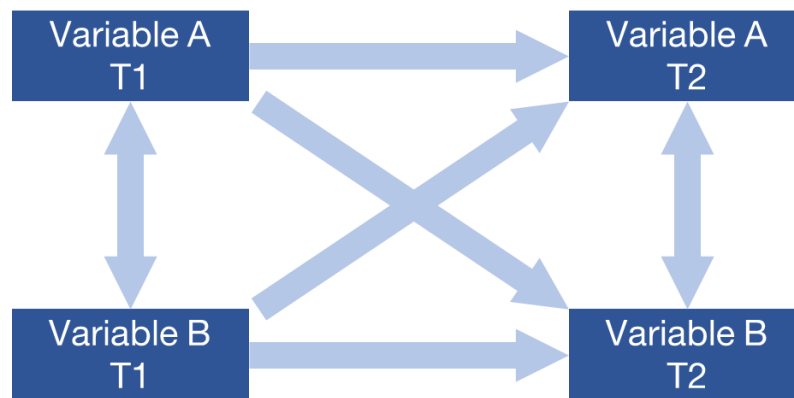


Figure 4: Bidirectional relationships. Disentangling the links between two correlated cognitive outcomes or risk factors which evolve in time (from variable A at T1 to variable B at T2, and vice-versa) requires the use of longitudinal data and more complex statistical models such as the random-intercept cross-lagged panel model.

Cognitive development is thus a dynamic and complex process, shaped by the influences and synergies of a large array of environmental, biological and genetic factors. The set of abilities thus formed constitutes the foundations of academic skills learning.

1.2 Acquiring academic skills: building blocks and determinants

In contrast to other cognitive abilities such as language or reasoning, literacy and numeracy need to be explicitly taught to children. However, the success of this instruction hinges on the child's initial cognitive skills, which are the building blocks of reading and mathematics (1.2.1), as well as on her socio-emotional skills, which may affect her approach to learning and behavior in the classroom (1.2.2). In addition, other socio-demographic factors come into play, likely exerting an influence on academic skills learning on top of their early influence on early cognitive development (1.2.3). We now delve into how these three aspects govern the acquisition of academic skills.

1.2.1 The role of cognitive abilities

Both reading and mathematics are cumulative processes that build on some key cognitive abilities. We first present the links between general cognitive development (intelligence) and academic skills, before focusing on specific cognitive domains (language, visuospatial and motor development, and executive functions).

General cognitive ability, or intelligence

Intelligence tests were initially designed with the explicit purpose of predicting children's future educational success (Binet & Simon, 1904). Similarly, it is from the observation that multiple school examination scores were all positively correlated that Charles Spearman extracted the first measure of general intelligence (the 'g' factor) (Spearman, 1904). Therefore, it should come as no surprise that IQ is one of the best predictors of academic achievement – if not the best, depending on the outcome measure used. Thus, the correlation between intelligence test score and academic skills lies between 0.5 and 0.8 (Deary et al., 2007; Rohde & Thompson, 2007; Roth et al., 2015). Nowadays, several standardized tests have been developed by psychologists to measure human intelligence; the most widely used for children being the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) for those aged 3 to 7, and the Wechsler Intelligence Scale for Children (WISC) for those aged 6 to 16. These tests measure the principal cognitive functions of an individual: processing speed, working memory, verbal comprehension, fluid reasoning, and visual spatial skills, summed up in a total score: IQ (standardized with a mean of 100 and standard deviation of 15). The full-scale IQ score and the five subcomponents are thought to correspond, respectively, to the *g* factor and five broad abilities in the Cattell-Horn-Carroll (CHC) intelligence theory, in which human intelligence is modelled as a hierarchical structure with the *g* factor at the top stratum, hypothesized to be at the core of all broad abilities in the stratum beneath (Schneider & McGrew, 2012). General intellectual ability as measured by IQ tests can also be broken down into two more comprehensive components: verbal intelligence, and non-verbal intelligence, or crystallized (*gc*) and fluid intelligence (*gf*) (Cattell, 1963; Horn & Cattell, 1966). While both *gc* and *gf* are well correlated with academic performance, crystallized intelligence seems to have a higher predictive power than fluid intelligence – which makes sense, since crystallized intelligence encompasses acquired knowledge, reflecting prior learning

Chapter 1 - General introduction

(correlation of 0.36 to 0.65 for crystallized intelligence versus 0.26 to 0.40 for fluid intelligence; Postlethwaite, 2011).

We now look beyond these general standardized measures of cognition to understand how various components of cognitive ability support the acquisition of academic skills.

Language abilities

Language abilities are the backbone of learning to read, as one can easily imagine; but they are also essential in learning mathematics. Several aspects of language are crucial in the acquisition of literacy. The first one is phonological processing, which is the ability to perceive, store, access and manipulate speech sounds. A particularly useful component of phonological processing is phonological awareness (being aware of and manipulating speech sounds), which enables children to map graphic symbols to the sounds of spoken words (at a sublexical level), and hence plays an important role in decoding and spelling. Phonological awareness is the best predictor of word recognition (Melby-Lervåg et al., 2012) and a good predictor of spelling (Landerl & Wimmer, 2008; Lervåg & Hulme, 2010). The second aspect of early language ability that is crucial in literacy acquisition is language comprehension. This includes vocabulary (mapping phonological representations onto semantic representations), which is essential for reading comprehension (Hjetland et al., 2020; Ouellette, 2006). Beyond vocabulary, grammar (the implicit knowledge of syntax and morphology) (Durand et al., 2013; Hjetland, 2018; Hjetland et al., 2020; Lehl et al., 2020; Muter et al., 2004; NICHD Early Child Care Research Network, 2005; Su et al., 2017) and conceptual knowledge (the understanding of concepts and classifications) (Hjetland et al., 2020; National Early Literacy Panel, 2008; Storch & Whitehurst, 2002) play an important role in reading comprehension.

In parallel, language abilities also play multiple roles in the development of numeracy skills. Indeed, children need to associate the rote-learned number words with the quantities they represent (Geary, 2013). Besides, simple arithmetic facts such as multiplications seem to be stored and retrieved from long-term verbal memory (Dehaene & Cohen, 1995). Lastly, in order to solve an arithmetic problem presented in sentences, children need to use their vocabulary and language comprehension abilities to understand the problem and translate it into an equation (Fuchs et al., 2010), and often keep the elements of the problem in verbal working memory. Thus, language skills have been found

to predict arithmetic abilities as well (Durand et al., 2005; Fuchs et al., 2010; Träff et al., 2018; Zhang et al., 2017).

Visuospatial abilities

Visuospatial abilities are important in the acquisition of both reading and mathematics. On one side, visuospatial skills are necessary to identify letters and segment written words into graphemes (letter or combination of letters transcribing phonemes). Few studies have examined the role of visuospatial skills in non-pathological reading, but the National Early Literacy Panel (2008) reported low univariate correlations with reading comprehension and word identification (around 0.2). In particular, deficits in visual attention have been proposed to account for the occurrence of developmental dyslexia in some children (Facoetti et al., 2010; Vidyasagar & Pammer, 2010). However visuospatial impairments could be a consequence rather than a cause of reading disorders (Ramus, 2003), and hence not be an early predictor.

On the other side, visuospatial abilities are an important foundation of numeracy acquisition. Indeed, children's arithmetic abilities partly lie on the development of an accurate linear mental representation of quantity (Siegler & Booth, 2004). In addition, spatial processing helps to solve complex arithmetic problems which require multistep calculations (Dehaene & Cohen, 1995). Lastly, in arithmetic word problems³, visuospatial abilities may support the construction of a visual schematic representation of the problem, which in turn may improve performance (Boonen et al., 2013). Thus, visuospatial abilities have been found to be correlated with higher results in arithmetic concurrently (Hawes et al., 2019; Reuhkala, 2001; Träff et al., 2018) and longitudinally (Yang et al., 2019; Zhang et al., 2014, 2017).

Fine motor abilities

Motor development is an important area of cognitive development in the first years of life, which has been purported to foster the acquisition of academic skills in various ways. In particular, fine motor skills – “small muscle movements that require close eye–

³ E.g., “A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?”

hand coordination” (Luo et al., 2007a) – may benefit both reading and mathematics because children with better fine motor skills have more opportunities to engage in learning activities promoting academic success (Suggate et al., 2019). Beyond this, fine motor skills may foster mathematic skills due to the fact that better finger-based representations of magnitudes may support the development of number sense – indeed, finger gnosis predicts later numerical abilities (Costa et al., 2011; Noël, 2005; Penner-Wilger et al., 2007). Thus, fine motor skills have been found to be positively associated with both reading outcomes (Cameron et al., 2012; Grissmer et al., 2010a; Pitchford et al., 2016) and mathematic outcomes (Carlson et al., 2013; Gashaj et al., 2019; Gomez et al., 2015; Grissmer et al., 2010b; Hawes et al., 2019; Luo et al., 2007b).

Executive functions

Executive functions are a set of higher-order cognitive processes which allow us to plan, regulate and monitor goal-directed behavior. They involve the following key functions: inhibition, selective attention, working memory and cognitive flexibility. These processes are important in the acquisition of academic skills, since they enable children to regulate, control and manage their learning. When reading, children have to select and hold in memory only the information which is useful to understand the meaning of the text – which implies good inhibition and working memory skills. Similarly, when solving an arithmetic problem, the child needs to select the relevant information and hold and manipulate numbers in their mind. Thus, inhibition and working memory have both been associated with better performance in both reading (Borella et al., 2010; Butterfuss & Kendeou, 2018; Chiappe et al., 2000; Seigneuric & Ehrlich, 2005) and mathematics (Alloway & Passolunghi, 2011; Holmes & Adams, 2006; Yang et al., 2019; Zhang et al., 2017). Attention is also of prime importance to keep children engaged in learning endeavors and is a predictor of both mathematics and reading achievement (Duncan et al., 2007; Rhoades et al., 2011; Steele et al., 2012).

Given their role in regulating and monitoring goal-directed behavior, executive functions exert a crucial role in socio-emotional development⁴, which is also an important predictor of academic achievement.

⁴ In fact, many of the cognitive processes included in executive functions seem to have their parallel in socio-emotional skills (e.g., inhibiting impulsive behaviors) (Riggs et al., 2006). Drawing the line between cognitive abilities and social, emotional and behavioral skills may thus seem artificial; however, given that the latter

1.2.2 The role of social, behavioral and emotional skills

Binet and Simon, the creators of the first broad test of intelligence, were aware of its limitations and acknowledged that success in school does not only depend on cognitive ability, but also on skills such as the capacity to maintain attention, grit or conscientiousness: “a child, even if intelligent, will learn little in class if he never listens, if he spends his time in playing tricks, in giggling, in playing truant” (Binet & Simon, 1916). A lot of emphasis has been put on the importance of such skills over the past decade in the context of education, under the umbrella terms of ‘socio-emotional skills’, ‘soft skills’, ‘personality traits’ or ‘non-cognitive skills’⁵. They encompass various aspects of behavioral and emotional regulation, self-perception, and social interactions, and have been associated with academic achievement.

The Big Five personality factors

One of the main framework for assessing socio-emotional skills is the Big Five model, measuring the following characteristics: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism/emotional stability (OCEAN). These characteristics have been found to predict academic achievement independently from general cognitive ability (Borghans et al., 2016). In particular, conscientiousness, and to a lower extent, openness to experience and agreeableness, exert an important influence on academic results across school levels; while extraversion and emotional stability seem to matter solely during primary school years (Poropat, 2009). For example, self-discipline, an important aspect of conscientiousness, seems to foster academic skill learning through homework completion and appropriate behavior in the classroom (Duckworth et al., 2012; Duckworth & Seligman, 2005).

Mental health/behavioral problems

Another common approach to measuring socio-emotional skills is the psychopathology framework, measuring deficits in these skills. These deficits can be

have traditionally been studied separately as predictors of academic achievement, we also make the distinction throughout this dissertation.

⁵ Although this last term does not feel accurate given the inherently cognitive origin of behavioral and emotional regulation.

Chapter 1 - General introduction

classified in two broad categories: internalizing problems (e.g. symptoms of anxiety, depression, or social withdrawal) and externalizing problems (e.g. impulsivity, aggressive behaviors or hyperactivity). Behavioral/emotional disorders and personality factors have been shown to partly overlap (Muris et al., 2005; Santos & Primi, 2014), as some of the mental health problems are tightly related to the Big Five components. These behavioral and emotional difficulties have long been associated with academic achievement, through their effects on individual learning and behavior in the classroom. Children with externalizing problems may learn less due to inattention and misbehavior, while internalizing problems (e.g. anxiety, depression) may reduce the motivation, participation and concentration of pupils in class; and both are detrimental to peer interactions (Bub et al., 2007). Thus, longitudinal studies have shown associations between lower academic achievement and early socio-emotional problems (Breslau et al., 2009), and in particular externalizing problems (Masten et al., 2005; Moilanen et al., 2010; but see Duncan et al., 2007).

1.2.3 The role of socio-demographic factors

Given that a number of socioeconomic and demographic factors participate in shaping cognitive development, it is expected that these predictors are associated with later academic skills through their early influence of cognitive abilities and socio-emotional skills. However, it is likely that, on top of this early effect, environmental factors have a more direct influence on the acquisition of academic skills. We now review evidence and potential mechanisms underlying this direct pathway.

Socioeconomic factors

The family's socio-economic status (SES) (encompassing parental education, income and occupation) has long been associated to children's educational outcomes. Thus, children from higher SES backgrounds have better academic skills than their low-SES peers (Sirin, 2005), partly for reasons other than differences in baseline cognitive development. Indeed, a high SES can help a child succeed at school in a number of ways. Parents with higher SES are more likely to assist their child with their homework (Tam & Chan, 2009), provide the child with the adequate social capital – such as knowledge of the

school culture and norms, as well as better relationships with teachers (Sirin, 2005) –, but also bolster the child's academic ambitions (Guyon & Huillery, 2016). Higher-SES parents are also more likely to provide their child with private tutoring if needed (Bray, 2011), and place their child in schools with higher achieving pupils, thus triggering positive peer effects on achievement (Burke & Sass, 2013; Hanushek et al., 2003). Another way SES may affect the acquisition of academic skills is through the internalization of negative stereotypes by disadvantaged pupils (Nguyen & Ryan, 2008; Steele & Aronson, 1995), or even biases in grading by school teachers (Hanna & Linden, 2012; Sprietsma, 2013).

Sex

Girls have been shown to outperform boys in school marks and examinations in virtually all subjects (Deary et al., 2007; Voyer & Voyer, 2014). However, several meta-analyses studying differences in achievement test scores have shown that girls perform better at language tests (Hedges & Nowell, 1995; Hyde & Linn, 1988), while boys perform better in mathematics (Else-Quest et al., 2010; Hyde et al., 1990; Reilly et al., 2015), albeit not consistently (Lindberg et al., 2010). These findings have been replicated by international data from the PISA studies (Programme for International Student Assessment), with 15-year-old girls outperforming boys in reading assessments in most participating countries, and the other way around in mathematics (OECD, 2015). While differences in cognitive and socio-emotional abilities may partly explain these gaps (e.g. self-discipline, see Duckworth & Seligman, 2006), it is likely that other factors participate. Thus, the internalization of negative stereotypes that we previously mentioned for disadvantaged students may also be at play for girls, especially in mathematics (Nguyen & Ryan, 2008; Spencer et al., 1999). Teachers may also play a role in amplifying or reducing sex differences in achievement. For example, primary school female teacher's math anxiety may reinforce female pupils' stereotype regarding girls' lower abilities in mathematics, thereby undermining their performance (Beilock et al., 2010). On the other hand, the perception that girls are at a disadvantage may lead teachers to give them a boost when grading (Breda & Ly, 2015; Falch & Naper, 2013; Lavy, 2008; Protivínský & München, 2018; Terrier, 2015; but see Lafontaine & Monseur, 2009a).

The acquisition of academic skills is thus supported by a rich set of cognitive and socio-emotional skills – themselves shaped by a complex network of environmental, biological and genetic factors, as we have seen –, and fostered, or hampered, by socio-

demographic characteristics (summarized in **Figure 5**). In this dissertation, we bring together these separate strands of research to study how these different components together determine children's academic outcomes in France.

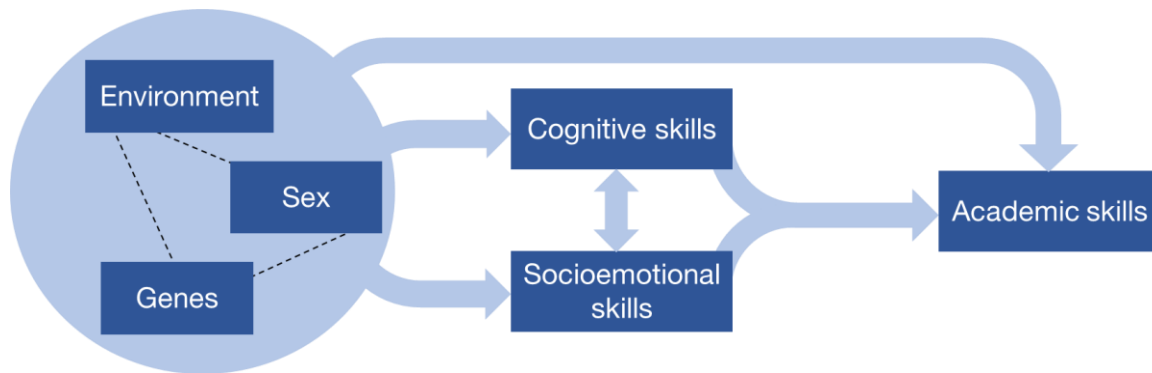


Figure 5: The acquisition of academic skills. A complex interplay of environmental, individual, cognitive and socioemotional factors influences a child's ability to learn.

1.3 Studying the acquisition of academic skills in France

The vast majority of studies that we cited so far concerned children raised in English-speaking countries, notably the US and the UK. Similar studies in different cultural and economic contexts are thus needed to generalize findings, and provide country-relevant information to national actors in the education sphere. In this research work, we focused on France, analyzing data from two longitudinal databases. In this section, we first present some context on academic achievement in France (1.3.1), before introducing the two cohorts on which this work is based on (1.3.2) and detailing the objectives and research questions of this dissertation (1.3.3).

1.3.1 Learning in France: an outlook

International comparisons have provided insightful knowledge of the state of academic skills learning in France. However, few French studies have focused on the multiple pathways which make a child a better learner.

Insights from international comparisons

The latest 2018 PISA study confirmed the observations from the previous PISA cycles, highlighting deep-seated challenges that France needs to tackle. While students in France scored slightly higher than the OECD average in both reading and mathematics (see **Figure 6**), more than 20% of French students still do not achieve a minimum level of proficiency in reading and mathematics by the end of middle school, thus impeding their educational prospects and functioning in daily life (OECD, 2019). More worrying, the gap between the highest- and lowest- achieving students in reading has widened over the years, with the top-achievers improving their performance, while that of the bottom-achievers' declined – and in mathematics, all students saw their performance decrease. Lastly, France remains one of the countries with the starkest gaps across socio-economic statuses. For instance, a low-SES student in France is 7 times more at-risk of low reading performance than a high-SES student – while on average across OECD countries, the odds are equal to 5 (OECD, 2019). Understanding how the different environmental and cognitive factors play in helping children acquire stronger academic skills is important to address these issues.

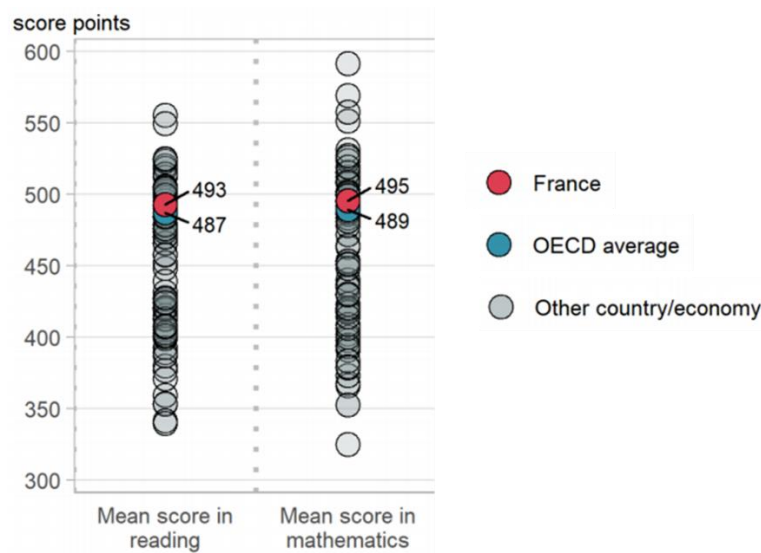


Figure 6: Performance at PISA 2018. Source: OECD (2019), *PISA country note – France*. URL: https://www.oecd.org/pisa/publications/PISA2018_CN_FRA.pdf

Predictors of academic achievement in France

Studies examining both the environmental and cognitive predictors of academic skills in France are very scarce, especially in large samples from the general population. Indeed, most studies focused on particular aspects of what shapes children's achievement. For example, Ben Ali & Vourc'h (2015) have looked at the social determinants of academic skills in middle school, highlighting the large social inequalities in achievement. However, they did not take into account socio-emotional or cognitive factors. To the best of our knowledge, the only study to have simultaneously assessed the environmental, cognitive and socio-emotional predictors of academic skills in France is that of Fluss et al. (2009), who looked at the effects of SES, behavioral and emotional problems, and cognitive skills, on reading ability (measured at the same time) among 1062 elementary school students in Paris. Their results emphasize large differences in SES, as well as influences of phonological awareness and attention deficits.

In the present work, we adopted this approach in order to understand the multidimensional factors which influence the acquisition of academic skills in France, through the longitudinal analysis of two cohort studies.

1.3.2 Data: two French cohort studies

This dissertation analyzes data from two French cohort studies which have the common advantage of gathering information on children's academic skills at about the same age (11 to 14 years old) as well as on their cognitive and socio-emotional abilities, and on a range of environmental factors: the EDEN mother child cohort, and the DEPP Panel 2007.

The EDEN mother-child cohort

The EDEN cohort (Heude et al., 2016) is an epidemiological birth cohort led by Inserm (*Institut national de la santé et de la recherche médicale*) research teams, which recruited pregnant women in the cities of Nancy and Poitiers, France. The study was approved by the Ethical Research Committee (Comité consultatif de protection des personnes dans la recherche biomédicale) of Bicêtre Hospital and by the Data Protection Authority (Commission Nationale de l'Informatique et des Libertés). Informed written consents were obtained from parents for themselves at the time of enrollment and for the newborn after delivery. Enrollment started in February 2003 in Poitiers and in September 2003 in Nancy and lasted for 27 months in each center. The sample at the time of inclusion consisted in 2,002 pregnant women seen during a prenatal visit at the departments of Obstetrics and Gynecology of the French University Hospitals of Nancy and Poitiers before their twenty-fourth week of amenorrhea, who agreed to participate and matched the inclusion criteria. Women with a personal history of diabetes, twin pregnancy, intention to deliver outside the university hospital or to move out of the study region within the following 3 years, and who could not speak French, were excluded from the sample. The participation rate among eligible women was 53%. 1,907 women out of 2,002 were still in the cohort at delivery.

Detailed data on children's environment and cognitive development were regularly collected from birth to 11.5 years old (age at the last wave: Mean=11.56, SD=0.51) through different means, including:

- **Psychometric tests of cognitive abilities at 3 and 5.5 years old:** at 3 and 5.5 years old, children's cognitive abilities were assessed at home by a trained psychologist, by means of a range of psychometric tests. In the dissertation, we only used cognitive data collected at 5.5 years old because they include a more complete set of cognitive abilities:

the full Wechsler Preschool and Primary Scale of Intelligence - III (WPPSI-III) (Wechsler, 1967, 2004), language tests from the NEUROPSYCHOLOGICAL assessment (NEPSY) battery (Kemp et al., 2001; Korkman et al., 2003), the Peg moving task (Nunes et al., 2008), and an early numeracy test, the Number Knowledge Test (NKT) (Okamoto & Case, 1996).

- **Clinical records at birth:** children's birth characteristics, including birthweight, gestational age and sex, were recorded in the hospital at delivery.

- **Parental questionnaires from pregnancy to 11.5 years old:** Mothers regularly answered questionnaires regarding their habits during pregnancy (e.g. alcohol consumption), parental characteristics (e.g. level of education or income), their interactions with their child, and own mental health.

- **Behavioral questionnaires at 3, 5.5, 8 and 11.5 years old:** at 3, 5.5, 8 and 11.5 years old, mothers answered the Strength and Difficulties Questionnaire (SDQ) (Goodman, 1997, 2001; Shojaei et al., 2009), which assesses behavioral and emotional problems in their child (namely, emotional symptoms, inattention/hyperactivity, conduct problems, peer relationship problems, and prosocial behavior). In addition, at 11.5 years old, mothers also completed the Child Behavior Checklist (CBCL) and children completed the Mental Health and Social Inadaptation Assessment for Adolescents (MIA) and the Strengths and Weaknesses of ADHD-symptom and Normal-behavior (SWAN). In this dissertation, we only used scores from the SDQ at 5.5 years old, alongside the psychometric scores at the same age.

- **Online academic skills tests at 11.5 years old:** at 11.5 years old, children were administered a range of tests measuring academic skills as well as broader cognitive abilities at home, on the family computer. In this dissertation, we used solely tests measuring academic skills, i.e. reading, spelling, and arithmetic.

The DEPP Panel 2007

The DEPP Panel 2007 (Trosseille et al., 2013) is a large cohort study led by the Direction de l'Évaluation, de la Prospective et de la Performance (DEPP; French Ministry of Education). It followed 34,986 French students from their first year of middle school in 2007 (grade 6, 11 years old) to their last year of middle school (grade 9, 14 years old). The study was compulsory and approved by the National Council for Statistical Information (CNIS) (visa n°2008A061ED and 2011A082ED), ensuring public interest and conformity with ethical, statistical and confidentiality standards. The sample was randomly selected

from an exhaustive sampling frame, ensuring representativeness by balancing available characteristics (region, public/private status of the school, urban unit, school establishment, age of entry in grade 6). The sample was constituted in such a way as to be representative of the population of French middle school students, with a slight over-representation of students in schools belonging to the *Réseau Ambition Réussite* (Success Ambition Network – schools in disadvantaged areas). Data was collected through questionnaires and tests at several time points from middle school to high school, including:

- **Standardized academic and cognitive skills tests at 11 and 14 years old:** in grade 6 and grade 9, the DEPP administered at-school standardized tests of academic skills to children, including tests of grammar, phonology, mathematic and school vocabulary, as well as a nonverbal reasoning test, the Chartier’s Reasoning Test on Playing Cards (*Raisonnement sur Cartes de Chartier*, RCC) (Terriot, 2014).
- **Conative skills questionnaires at 11 and 14 years old:** along with these academic skills tests, children had to answer self-efficacy and academic motivation questionnaires, derived from the Children’s Perceived Self-Efficacy scales (Bandura, 1990) and the Academic Self-Regulation Questionnaire (Ryan & Connell, 1989).
- **School marks and national examination results:** in grade 6 and grade 9, teachers’ marks were reported, as well as students’ results in national examinations (*Examinations nationales de 6e* and the *Diplôme National du Brevet*).
- **Parental questionnaires at 11 and 14 years old:** at both time points, parents filled in questionnaires informing us on the social, economic and cultural environment in which the child evolves (e.g. parental socio-economic status, type of school, parental involvement in school).
- **Follow-up educational outcomes in high school:** additional information directly recorded by the schools was reported, informing on the academic path taken by the child after middle school (e.g. vocational or general).

As can be seen, these two cohort studies each have their strengths and complement each other in a number of ways: the EDEN cohort boasts a rich set of psychometric tests and pre- and post-natal environmental factors, as well as an extended longitudinal design still unequalled in a French birth cohort (from pregnancy to 11 years old), while the DEPP cohort features a very large and representative sample (about 35,000 children), multiple measures of academic skills and of socio-economic characteristics. We took advantage of

Chapter 1 - General introduction

these different strengths to answer distinct research questions exploring the cognitive and environmental predictors of academic skills in France.

1.3.3 Objectives and research questions of this dissertation

This dissertation aims at providing a further understanding of what fosters the acquisition of academic skills in children. It has two leading threads: first, the necessity to portray and assess the multidimensional factors influencing academic outcomes – cognitive, socio-emotional, and environmental⁶; second, the need to study these in understudied contexts, here France. This thesis is organized in the following chapters:

Chapter 2 – Intelligence and academic skills

In this chapter, we examine the extent of the association between intelligence and academic skills in two studies. The first one assesses the strength of this relationship in the large, representative sample of students from the DEPP Panel 2007 as well as the socio-economic and conative predictors of academic skills and their progression, beyond the role of IQ. The second one inquires whether this relationship still holds in the high-IQ range and whether intellectually gifted students suffer from school failure, as often depicted in the media.

Chapter 3 – Early predictors of arithmetic skills

In this chapter, we dig into one component of academic skills: numeracy. In a first study, we assess the predictive role of a rich set of pre-school environmental, individual, cognitive and socio-emotional skills on arithmetic word problem solving at 11.5 years old in the EDEN mother-child cohort. We also estimate the extent to which the effects of early environmental and individual factors are mediated through pre-school cognitive and socio-emotional skills. In a second study, we break down arithmetic into three types of calculations – addition, subtraction and multiplication – in order to examine whether they are predicted differentially by early cognitive abilities.

⁶ The study of genetic predictors of academic achievement is beyond this research work, but will nevertheless be discussed in the general discussion (Chapter 6).

Chapter 4 – Early predictors of literacy skills

In this chapter, we focus on the second component of academic skills: literacy. We examine the role of a broad range of early cognitive, socio-emotional, individual and environmental variables on different literacy skills (reading comprehension, reading accuracy, reading speed, and spelling) at 11.5 years old in the EDEN cohort. In addition, we ask whether the effects of early environmental factors are mediated through early cognitive and socio-emotional abilities.

Chapter 5 – Sex differences in academic skills

In this chapter, we study the differences between girls and boys in literacy and mathematics, enlightening results from the previous chapters by examining the influence of evaluation characteristics on the gender gap. We do so by comparing boys' and girls' results at three different types of evaluations (teacher evaluations, national examinations, and standardized achievement tests) both in French and mathematics in the DEPP Panel 2007.

Chapter 6 – General discussion

In this last chapter, we synthesize, connect and discuss the original contributions of the dissertation, acknowledging its limitations and identifying practical implications and perspectives.

Chapter 2 - Intelligence and academic skills

In this chapter, we explore two aspects of the well-known association between intelligence and academic skills, answering the following questions: what's the strength of this association in a representative sample of French students; and what are the predictors of achievement and of its progression beyond intelligence (Part 2.1)? Does the positive relationship between IQ and achievement hold in the high-IQ range? Are intellectually gifted students really more at risk of school failure (Part 2.2)?

2.1 Predictors of the IQ-achievement gap

This section is based on the following published article:

Guez, A., Panaïotis, T., Peyre, H., & Ramus, F. (2018). Predictors of the IQ-achievement gap in France: A longitudinal analysis. *Intelligence*, 69, 104–116.

2.1.1 Abstract

Why do some children under-perform at school relative to their level of cognitive abilities? So far, previous studies on the topic have been conducted on cross-sectional data or have focused on a limited range of predictors. In this large longitudinal study on 23,258 French middle school students, we examined the relative effects of a wide range of contextual and individual factors on academic performance beyond the effect of non-verbal IQ. Data were analyzed using a structural equation modeling approach, cross-sectionally and longitudinally. Cross-sectional models revealed that self-efficacy, school environment, parental education and sex were the most predictive factors of achievement independently from non-verbal IQ (the latter being by far the best predictor). A longitudinal analysis showed that school environment and parental education also significantly affected progression between grade 6 (11-12 years old) and grade 9 (14-15 years old), while non-verbal IQ and other factors played a minor role.

Keywords: academic achievement, intelligence, socio-economic environment, self-efficacy, sex, longitudinal study

2.1.2 Introduction

The first tests measuring cognitive abilities were initially developed in order to predict children's future educational outcomes (Binet & Simon, 1904). Likewise, Charles Spearman extracted for the first time a measurement of general intelligence (the 'g' factor) after observing the positive correlations that exist among school examination scores (Spearman, 1904). Measurements of general intelligence are indeed good predictors of academic achievement, such that their correlation lies between 0.5 and 0.8 (Deary et al., 2007; Rohde & Thompson, 2007; Roth et al., 2015). However, a substantial part of the variance in academic achievement remains unexplained by general intelligence. Thus, some children under- or over-perform at school relative to what is expected from their level of cognitive ability. This discrepancy between an individual's actual and expected level of academic performance given IQ is known as the IQ-achievement gap (Flynn, 1991; Gordon, 1976). Understanding the different determinants of this gap is of great importance not only for researchers but also for education professionals and policy-makers. Nevertheless, most studies have focused on only one or two specific predictors of achievement beyond intelligence, without considering other important factors.

A first strand of research has looked at the effect of the socio-economic environment on achievement. Family socio-economic status (SES) is a multi-faceted concept (mainly involving income, occupation, and education) that may affect academic performance for different reasons. First, higher financial resources may allow parents to afford private lessons, as well as cultural and educational objects, trips and leisure, which can help children succeed in school. Higher educated parents are also more capable to help their child with homework and provide them with appropriate interactions. Second, SES provides social capital that may help succeed in school (knowledge of the culture and norms valued in school, better relationships with teachers) (Sirin, 2005) as well as higher ambition (Guyon & Huillery, 2016). Numerous studies have demonstrated the predictive role of socio-economic status (SES) (see the meta-analysis by Sirin, 2005) and parental involvement in school (Cheung & Pomerantz, 2012; Hill & Craft, 2003) on academic achievement, but rare are those which took into account measurements of cognitive ability. Not including IQ as a predictor could be misleading because the association between socio-economic status and academic achievement is known to be partly accounted for by general intelligence and mediated by genetic factors (Krapohl & Plomin, 2016). Most of the studies that did include IQ as a covariate had a small sample size or were cross-sectional (Bacete

& Remírez, 2001; Johnson et al., 2007; Jurecska et al., 2012). More recently, a longitudinal analysis of a large sample of UK students showed that SES had an effect on achievement and its progression from age 7 to 16 independently from IQ (von Stumm, 2017).

Other studies have investigated the influence of self-efficacy, namely individuals' beliefs in their own ability to achieve intended results. The idea that self-efficacy affects performance stems from early works by Bandura and colleagues, which argued that it has a positive motivational influence on achievement: students with higher self-efficacy set more difficult goals, work harder, and are more persistent when they face difficulties (for a review, see (Zimmerman, 2000). Again, several studies have looked at the effect of self-efficacy on future academic performance (Talsma et al., 2018), but few of them included intelligence as a predictor. Given the (partly genetic) overlap between self-efficacy and cognitive abilities (Greven et al., 2009), it is important to control for IQ when estimating the link between self-efficacy and achievement. Among the few studies that addressed this issue, Chamorro-Premuzic, Harlaar, Greven, & Plomin (2010) found that self-efficacy measured at age 9 increased achievement at age 12 independently from past levels of achievement and cognitive abilities. Similarly, Kriegbaum, Jansen, & Spinath (2015) have shown that self-efficacy was the strongest motivational predictor of performance at PISA mathematics tests at age 16 beyond prior achievement and intelligence (measured at age 15).

Lastly, many researchers have examined the effect of sex, showing the existence of sex differences in academic achievement that cannot be explained by intelligence. For example, Deary et al. (2007) have shown that girls scored higher than boys in almost all subjects of a national exam in the UK (effect size for the overall score was Cohen's $d=0.3$), even when cognitive ability was taken into account. General intelligence accounted for 49.2% of the variance in GCSE total score, while sex contributed to 3.2% of its variance. Likewise, although in smaller samples, Duckworth & Seligman (2006) and Steinmayr & Spinath (2008) have found an effect of sex on academic performance independent from IQ, such that girls performed higher. Such a positive effect may be the result of girls showing higher levels of self-discipline (Duckworth & Seligman, 2006). On the other hand, other studies have pointed towards an underperformance of girls: for example, results from the PISA studies have shown that girls tend to perform lower than boys in some subjects, in particular in mathematics (OECD, 2015). This might be the consequence of lower levels of self-efficacy in mathematics: controlling for their performance level, girls tend to feel more anxious towards mathematics and have less confidence in their skills than boys (*idem.*).

Lower self-esteem and achievement, in particular in mathematics, might result from the internalization of negative stereotypes among girls (Spencer et al., 1999).

While these studies inform us about the predictive role of the socio-economic context, self-efficacy and sex on achievement and its progression beyond IQ, they do not provide information on the relative importance of the specific predictors studied compared with each other. Studies that did include a larger range of predictors also had a number of limitations. In a recent article, (O'Connell, 2018) examined the relative contributions of cognitive ability, personality, sex, family background and school characteristics to academic achievement. However, in that study, cognitive ability and personality tests were taken at age 13, while academic achievement was assessed at age 9, which considerably hinders the interpretation of the results.

In order to overcome the above limitations, the present study draws on a rich longitudinal database of French middle school students, which allowed us to estimate the relative contributions of intelligence, self-efficacy, sex and various indices of the socio-economic environment to academic achievement, both cross-sectionally and longitudinally. Data on academic achievement, non-verbal intelligence, individual and contextual factors were all collected twice: in grade 6 (11-12 years old) and grade 9 (14-15 years old). Our study design thus enabled us to estimate the effect of all predictors on academic achievement cross-sectionally, and on the progression in academic achievement between grade 6 and grade 9.

Students' fluid intelligence was assessed with a non-verbal reasoning test (Chartier's Reasoning Test on Playing Cards). We refer to this score as 'non-verbal IQ' throughout the paper. The term 'IQ-achievement gap', left unchanged for the sake of simplicity, thus refers in this paper to the gap between actual achievement and achievement predicted by non-verbal IQ.

We focused on middle-school because international comparisons consistently rank France very low at this particular level: the PISA studies have repeatedly pointed out the stark inequalities at school that face middle-school students from disadvantaged social classes in France (OECD, 2016b). Besides, it is worth noting that most studies on academic achievement have been conducted in the US or UK, and very few in France. In one of the rare French longitudinal studies, (Ben Ali & Vourc'h, 2015) have shown a strong impact of social and familial environment on academic progression, but without taking cognitive abilities into account. To the best of our knowledge, no study has examined the predictors

of the IQ-achievement gap in France – an analysis that would provide French policy-makers and actors in the education sphere with relevant information.

Based on the results of previous studies, we hypothesized that 1) the gap between achievement and non-verbal IQ is affected by 1a) the social, economic and cultural environment (i.e. the school environment, parental involvement in school, parental education, household income, cultural resources available and extracurricular activities), 1b) self-efficacy, 1c) sex. We further hypothesized that 2) the same factors would further amplify the gap during middle school (from grade 6 to 9).

2.1.3 Method

Sample

As part of a large study led by the *Direction de l'Évaluation, de la Prospective et de la Performance* (DEPP), French Ministry of Education, 34,986 children were followed from their entrance in first year of French middle school (grade 6 – mean age = 11.09, SD = 0.42) in 2007 to their last year of middle school (grade 9 – mean age = 14.21, SD = 0.52) (Trosseille et al., 2013).⁷ The study was compulsory and approved by the National Council for Statistical Information (CNIS) (visa n°2008A061ED and 2011A082ED), ensuring public interest and conformity with ethical, statistical and confidentiality standards. The sampling strategy consisted in randomly selecting a balanced sample from the characteristics observed in an exhaustive baseline survey. The sample was constituted in such a way as to be representative of the French population of middle school students with a slight over-representation of students in schools belonging to the *Réseau Ambition Réussite* (Success Ambition Network – schools in disadvantaged areas). We excluded from our working sample participants for whom the intelligence, academic achievement scores were missing or equal to zero, or for whom no index of socio-economic status was available. 23,258 participants were thus included in the present study (see **Figure A1**).

⁷ Panel d'élèves du second degré, recrutement 2007 - 2007-2013, DEPP - Ministère de l'Éducation [producteur], ADISP-CMH [diffuseur]

Measurements

In grade 6 and grade 9, students completed standardized tests measuring their level of academic performance, non-verbal intelligence and perceived self-efficacy. All tests were administered collectively, in paper/pencil format. In addition, parents were asked to fill a questionnaire giving information on the socio-economic environment and their own degree of involvement in their child's schooling.

Academic achievement. A battery of five tests measuring school-related skills was administered to students.

- *Phonics skills:* 'Odd-one-out' exercise where participants had to tick the word that does not have a common sound with the others (10 lists of 5 words each).

Example: *fer, aimer, verre, amer, hiver* (\fɛʁ, \e.me, \vɛʁ, \a.mɛʁ, \i.vɛʁ)

- *Grammar:* Cloze test (blank-filling task) composed of three short texts with missing logical connectors, determiners, or pronouns (20 items) (J. Aubret et al., 2006).

Example: « *Septembre ! C'est le mois.....choisit l'hirondelle pour partir vers le sud du Sahara.....elle peut passer l'hiver au chaud* ». ("September! This is the month.....the swallow chooses to fly towards the South of Sahara....it can spend its winter warm").

- *Mathematics:* Exercises of logic, mental arithmetic, problem solving, units and time calculations, and geometry (45 items). Questions were open-ended or multiple choice (F. Aubret & Blanchard, 1992; Blanchard & Berger, 1994; OECD, 2011).

Examples:

- $27 \times 20 = \dots$
- *Zoé est plus petite que Joëlle, et elle est plus grande que Cécile. La fille la plus grande s'appelle : 1) Cécile 2) Zoé 3) Joëlle* (Zoé is shorter than Joëlle, and she is taller than Cécile. The tallest girl is named: 1) Cécile 2) Zoé 3) Joëlle)
- *Reading comprehension:* Silent reading of three short texts, each of them followed by five questions. The task was completed in a limited time (12 min) (Aubret et al., 2006).
Example: *D'après le texte, combien de garçons escaladent le mur ?...* (According to the text, how many boys are climbing on the wall?....)
Quel est celui qui suit le sentier ? C'est... (Who is following the path? It is...)
- *Academic knowledge:* Multiple choice questionnaire where participants had to tick the best word associated with school manual words in the following disciplines: French, Mathematics, History, Geography, and Sciences (8 items each) (Lieury, 1996, 2012).

Example: *Batracien* (Amphibian)

- 1) *Grenouille* (Frog)
- 2) *Chimpanzé* (Chimpanzee)
- 3) *Oiseau de proie* (Bird of prey)
- 4) *Je ne sais pas* (I do not know)

The tests administered in grades 6 and 9 were very similar, except for some items that were changed in grade 9 to match students' higher level. Internal consistency was good or acceptable for all measurements (see **Table A1**). The relatively high correlations between tests in grade 6 and grade 9 suggest a good test-retest reliability for the different tests (see **Table 1**). 105 participants scored zero at all tests in grade 6, suggesting that they were not engaged in the task; hence they were excluded from the present study (as indicated in **Figure A1**). We ran an exploratory factor analysis with the five achievement tests in grades 6 and 9, which yielded a clear one factor solution in both grades (only one factor had an eigenvalue higher than 1, which explained 55% of the variance in grade 6, and 59% in grade 9). We used factor scores resulting from these analyses as our score of academic achievement in each grade (see **Table A2**).

Non-verbal IQ. Students' non-verbal IQ was assessed in grades 6 and 9 using the *Raisonnement sur Cartes de Chartier* test (RCC, Chartier's Reasoning Test on Playing Cards), that was designed to capture fluid intelligence (*gf*) (Chartier, 2012; Terriot, 2014). The test is made of 30 items assessing children's non-verbal logical reasoning skills, inspired from Raven's progressive matrices (Raven, 1998) but using playing cards. Each item is solved by determining which card (from a deck of 40 playing cards – ten of each suit) would fill the blank in an array composed of 4 to 12 cards. The RCC is scored as the number of items correctly completed in a limited time (20 minutes). Internal consistency was good ($\alpha = 0.88$ in grade 6 and 0.87 in grade 9; **Table A1**), and the correlation between RCC scores in grade 6 and RCC scores in grade 9 was relatively strong ($r = 0.61$), indicating a good reliability (**Table 1**). This correlation is also close to that found by Ramsden et al. (2011) in adolescents of similar age between the non-verbal scores from the Wechsler Intelligence Scale for Children (WISC-III) at time 1 (2004) and the Wechsler Adult Intelligence Scale (WAIS-III) at time 2 (2007/2008) ($r = 0.59$). The distribution of RCC scores was slightly negatively skewed, with a peak at zero (more marked in grade 6). We removed participants who scored zero from our analysis (as indicated in **Figure A1**), assuming they were not engaged in the task.

Table 1*Correlation Coefficients Between Tests in Grade 6 and Grade 9*

	Tests in grade 6						Tests in grade 9					
	1	2	3	4	5	6	7	8	9	10	11	12
Tests in grade 6												
1. Non-verbal IQ	1											
2. Phonics skills	0.37	1										
3. Grammar	0.46	0.50	1									
4. Mathematics	0.60	0.48	0.64	1								
5. Reading comprehension	0.43	0.43	0.61	0.57	1							
6. Academic knowledge	0.39	0.40	0.61	0.59	0.54	1						
Tests in grade 9												
7. Non-verbal IQ	0.61	0.33	0.44	0.59	0.39	0.38	1					
8. Phonics skills	0.36	0.51	0.48	0.46	0.40	0.40	0.39	1				
9. Grammar	0.46	0.46	0.73	0.63	0.57	0.59	0.51	0.53	1			
10. Mathematics	0.57	0.42	0.58	0.83	0.48	0.57	0.65	0.48	0.65	1		
11. Reading comprehension	0.43	0.39	0.62	0.57	0.56	0.56	0.48	0.43	0.67	0.59	1	
12. Academic knowledge	0.41	0.39	0.62	0.62	0.53	0.75	0.46	0.45	0.67	0.67	0.64	1

Note: Pearson correlation coefficients are reported. All coefficients are significant ($p < .0001$).

Correlations between the same tests administered in grade 6 and grade 9 are marked in bold.

Self-efficacy. In grades 6 and 9, students answered questions from the Children's Perceived Self-Efficacy scales (Bandura, 1990), translated into French. It is a 37-item questionnaire from which factors representing perceived *academic self-efficacy*, *social self-efficacy* and *self-regulatory efficacy* were extracted. The perceived *academic self-efficacy* score was constructed from 19 items that measure students' perceived ability to manage their learning, to master different academic subjects (mathematics, sciences, etc...), and to fulfill parents' and teachers' expectations. The perceived *social self-efficacy* score includes 13 items measuring efficacy regarding leisure group activities, the ability to form and maintain social relationships and manage interpersonal conflicts, and self-assertiveness. Lastly, the perceived *self-regulatory efficacy* score consists of 5 items measuring students' perceived ability to resist peer pressure to engage in high-risk activities (alcohol, drugs, transgressive behaviors). For each item, students had to evaluate their ability to perform each activity using a 5-points Likert scale. Internal consistency was good for the three indicators (Cronbach's α above 0.80). We ran an exploratory factor analysis with the three

self-efficacy indicators in grades 6 and 9, which yielded a clear one factor solution in both grades (only one factor had an eigenvalue higher than 1, which explained 40% of the variance in grade 6, and 35% in grade 9). We used factor scores resulting from these analyses as our score of self-efficacy in each grade (see **Table A3**).

Socio-economic and cultural environment. Parents or legal guardians filled a questionnaire in grade 6 and grade 9 evaluating the socio-economic and cultural environment in which the child evolves. We created eleven indicators.

Parental education. Parents reported their highest diploma, which we converted into years of education completed by each of them (from 0 to 18.5 years –18.5 years corresponding to a graduate degree; $M = 12$, $SD = 3.4$). We then took the mean of both parents as an indicator of parental level of education. When one parents' education was missing, only the other parents' education was taken into account. Mothers' mean education was slightly higher than fathers' (mean difference = 0.1405, $p < 0.0001$).

Household monthly income. Parents filled in household monthly income in grade 6 and grade 9. The relationship between achievement and income was non-linear: the higher the income, the lower the increase in achievement, i.e. the slope of the curve representing achievement as a function of income was positive but less and less steep as income increased (diminishing returns). Therefore we took the natural logarithm of income in order to correctly model this relationship.

Cultural objects in the house. Parents were asked in grade 6 whether there were books in the household, and if so how many. Their answer to the second question was coded in 4 categories: “between 1 and 29”, “between 30 and 99”, “between 100 and 199”, “200 or more”. The same was asked for CDs. We created a 5-category variable, adding a category “no books” to the four categories described above, which we then standardized; and did the same for CDs. The variable ‘cultural objects in the house’ is the mean of these two standardized variables (whose correlation was equal to 0.59).

Extracurricular activities. Parents reported in grade 6 and 9 whether their child was enrolled in: a sports club; a library; a music school; a theatre class; a scout movement (coded as 1 = “Yes”, and 0 = “No”). We took the mean of those answers as an indicator of extracurricular activities.

Parental involvement. We ran a principal component analysis of 13 variables linked to parental involvement in grade 6 and grade 9 (with promax rotation since we assumed the

factors would be correlated), which yielded five factors based on the scree plot (see **Table A4**). From these five factors, we created five scores measuring different aspects of parental involvement:

Frequency of conversations between parents and child about school. Parents reported how frequently they had conversations with their child about: homework, school learning, classmates, school life, and teachers. For each question there could be three different answers: “Never or almost never”, “From time to time”, and “Regularly”. We created a score of conversations about school by taking the mean of those 5 questions.

Frequency of conversations between parents and child about the future. Parents reported how frequently they had conversations with their child about their academic future and professional future. We created a score of conversations about the future by taking the mean of those 2 answers.

Involvement in school life. We created a variable measuring parents’ degree of involvement in the school by taking the mean of two dichotomous variables indicating whether parents belonged to a parents’ association, and whether they were class representatives.

Meeting with teachers. Parents were asked whether they had individual meetings with teachers on their own initiative, and on teachers’ initiative. Such meetings may originate from difficulties encountered at school by their child (such as a disrupting behavior or relational problems with classmates for examples). Following the results of the PCA, we created a variable for meetings with teachers taking the mean of those two dichotomous variables.

Commitment to help the child. This last variable is the mean of three dichotomous variables: one indicating whether parents went to a parents-teachers meeting, another indicating whether the child received private tutoring, and a last one indicating whether the child received help with homework. Parents-teachers meetings are meetings where all parents of the same class are invited at the same time to meet the class’ teachers, be informed on the class’ situation and their child’s schooling if they wish. The motivation to go to such meetings is thus very different from going to an individual meeting with a teacher.

School environment. In France, two structures for priority education schools existed at the time of the cohort: the RAR (*Réseau Ambition Réussite*, Ambition and Success Network) and the RRS (*Réseau Réussite Scolaire*, Academic Success Network). The RAR included the schools facing the greatest difficulties due to the local socio-

Chapter 2 - Intelligence and academic skills

economic environment (the 249 most disadvantaged schools, based on social and academic criteria), and the RRS (940 schools) those facing difficulties but not meeting all the criteria to be included in the RAR, and receiving extra-resources in order to counter school failure (Garrouste & Prost, 2015). We created a dummy variable indicating whether students attended a priority education school (RAR or RRS).

Parents' origins. We created a dichotomous variable equal to 1 if both parents were born abroad, and to 0 if at least one of the parents was born in France. We chose to group participants who had only one parent born in France with those whose parents were both born in France because they did not differ much in terms of achievement and non-verbal IQ, while those with both parents born abroad had significantly lower achievement and non-verbal IQ (for example, the mean Chartier's Reasoning Test on Playing Cards score for students with two parents born abroad was 12.9, versus 15.1 and 15.9 for students with respectively one and two parents born in France). Such lower results are likely to be the consequence of disadvantages linked to having both parents born abroad, such as speaking a language other than French at home, or a lack of knowledge regarding the French schooling culture.

Statistical Analysis: Structural Equation Models

We analyzed the data using structural equation models (SEM) (Bollen, 1989). Structural equation modeling is a statistical analysis technique that combines multiple regression with factor analysis. It allows us to construct latent variables from several observed variables – thus reducing measurement errors – and to model relationships between these latent variables and predictor variables. We used SEM with maximum likelihood estimation to assess the effect of our various predictor variables (measured, or latent in the case of self-efficacy) on our latent achievement outcome variables in grade 6 and 9. We used multiple imputation to handle the missing data in our predictor variables (for each analysis, ten imputed datasets were created, the analysis was performed on each of the imputed dataset, and parameters from each imputed dataset were averaged out to give the final parameters). We used standardized estimates since they are less affected by the scale of measurement and allow us to compare the relative influence of different predictors. We performed the analyses using SAS 9.4 for data cleaning, recoding and creation of variables as well as for descriptive statistics, and Mplus 8 to run SEMs. Statistical significance of parameter estimates was assessed using a two-sided test,

correcting for multiple testing with a Bonferroni correction (for a better readability, we reported only significance for $\alpha = 0.001$, which is roughly 0.05 divided by the number of tests – 28 in the cross-sectional models, 29 in the cross-sectional models, 29 in the longitudinal model). We evaluated goodness of fit using the comparative fit index (CFI), and the root mean squared error of approximation (RMSEA). Modification indices (Chi-square tests) were checked to test if there were significant correlations with any residual.

Cross-sectional models

We first assessed the relative predictive power of non-verbal IQ, self-efficacy, socio-economic index and sex on academic achievement in grades 6 and 9. We ran a first set of models without non-verbal IQ as a covariate (Models M6 and M9), and a second set including it in order to know how much of the effect of each predictor is explained by non-verbal IQ (Models M'6 and M'9). We hence estimated the following models at t equal to grade 6 and grade 9:

$$\text{Model M} : Ach_{it} = \beta_1 SE_{it} + \beta_2 Context_{it} + \beta_3 Sex_i + \varepsilon_t$$

$$\text{Model M}' : Ach_{it} = \alpha IQ_{it} + \beta_1 SE_{it} + \beta_2 Context_{it} + \beta_3 Sex_i + \varepsilon_t$$

Where Ach_{it} is a latent variable for general academic performance of student i at time t such that $Testscore_{it} = \lambda Ach_{it} + \xi_t$, with $Testscore$ a vector of the test scores at the five different achievement tests of student i at time t ; IQ_{it} is the non-verbal IQ score from the RCC test of student i at time t ; SE_{it} is a latent variable for self-efficacy of student i at time t such that $SE_score_{it} = \lambda SE_{it} + \xi_t$, with SE_score a vector of the three self-efficacy scores of student i at time t ; $Context_{it}$ is a vector of the socio-economic index of student i at time t ; and Sex_i is a sex dummy variable taking value 1 if student i is a female.

Longitudinal model

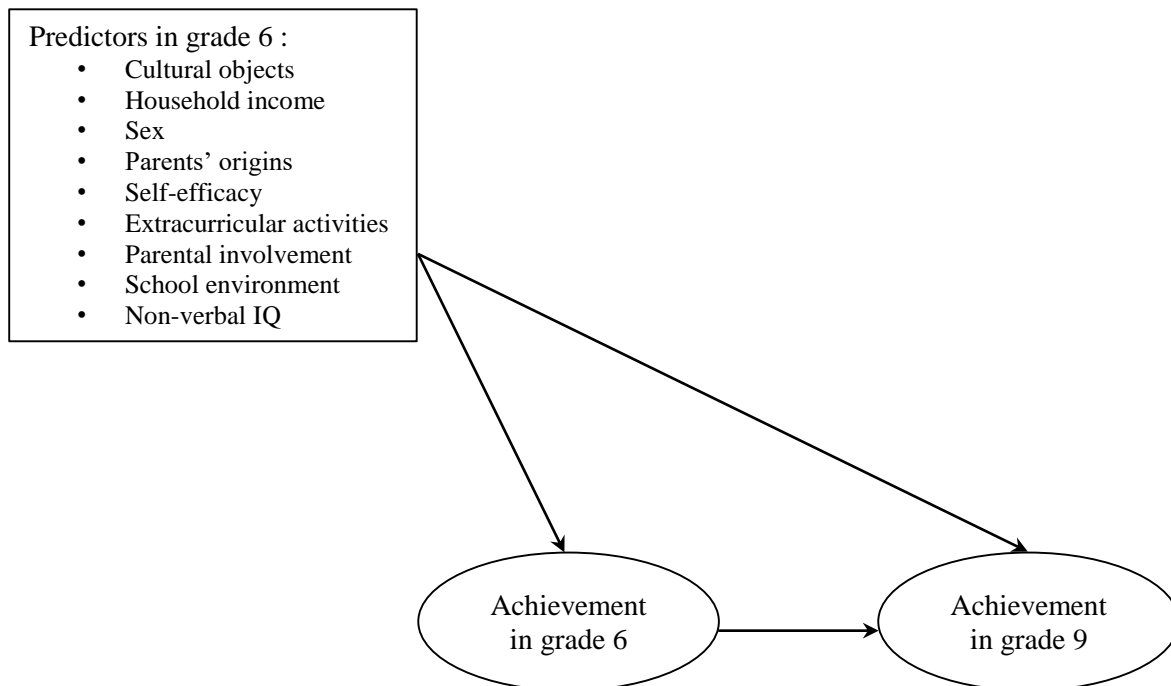
Next, we estimated to what extent the different factors affect students' academic progression during middle school. We performed a longitudinal SEM estimating the effects of predictors in grade 6 on achievement in grade 9, controlling for achievement in grade 6. We thus estimated the following set of equations with t =grade 9 and $t-1$ =grade 6:

$$\begin{cases} Ach_{it} = \theta Ach_{it-1} + \alpha IQ_{it-1} + \beta_1 SE_{it-1} + \beta_2 Context_{it-1} + \beta_3 Sex_i + \varepsilon_t \\ Ach_{it-1} = a IQ_{it-1} + b_1 SE_{it-1} + b_2 Context_{it-1} + b_3 Sex_i + \varepsilon_t \end{cases}$$

Figure 1 illustrates the model.

Figure 1

Longitudinal Structural Equation Model: Effects of Predictors on Academic Progression Between Grades 6 and 9



Note: We allowed errors of the same measured variables across time to be correlated (for example, score in Mathematics in grade 6 and in grade 9). Source: MENESR DEPP

2.1.4 Results

Analyses were conducted on a sample of $N = 23,258$ with non-verbal IQ and school achievement scores available and different from zero, and with socio-economic index available, in grades 6 and 9 (**Figure A1**). **Table 2** presents descriptive statistics of the participants. There were few missing data on socio-economic environment indices (less than 2.5% for all indices, with the exception of extracurricular activities – 9.5% – and

household income – 43.6% in grade 6) and on self-efficacy scores (less than 13% in grade 6).

Table 2

Summary Statistics of Included Participants in Grades 6 and 9

Variable	Grade 6			Grade 9		
	<i>N</i>	<i>% or M</i>	<i>SD</i>	<i>N</i>	<i>% or M</i>	<i>SD</i>
<i>Child characteristics (%)</i>						
Two parents born abroad	22,544	9.00		22,544	9.00	
Sex (female)	23,258	51.00		23,258	51.00	
<i>Socio-economic status</i>						
Parents' education (years)	23,201	12.08	3.39	23,201	12.08	3.39
Monthly income (EUR)	13,110	2980.62	1982.02	13,562	3331.32	3937.18
Priority education school (%)	23,258	17.00		23,258	16.00	
<i>Perceived self-efficacy (factor score)</i>						
Perceived self-regulation	22,215	0.07	0.96	23,002	0.18	0.68
Perceived academic self-efficacy	20,328	0.07	0.96	22,338	-0.54	1.03
Perceived social self-efficacy	21,168	0.02	0.97	22,495	0.01	0.94
<i>Academic skills (correct items)</i>						
Phonics skills (out of 10)	23,258	6.88	2.13	23,165	7.71	1.75
Grammar (out of 20)	23,258	8.98	4.28	23,181	10.90	4.08
Mathematics (out of 45)	23,258	26.95	8.70	23,211	27.75	10.00
Reading comprehension (out of 15)	23,258	9.75	3.01	23,190	10.33	4.10
Academic knowledge (out of 48)	23,258	28.25	7.30	23,245	31.00	8.11
<i>Non-verbal reasoning test (out of 30)</i>	23,258	15.96	5.86	23,258	18.58	5.81
<i>Parental involvement (%)</i>						
Parents attended a parent-teacher meeting	23,190	92.00		21,417	92.00	
Parents met teacher on their own initiative	23,139	29.00		17,729	41.00	
Parents met teacher on the teacher's initiative	23,141	20.00		17,156	34.00	
Parents part of a parents' association	23,160	14.00		21,962	12.00	
Parents are class representatives	23,165	10.00		19,252	10.00	
Help provided to student at home	23,143	91.00		19,825	73.00	
Private tutoring	22,709	9.00		21,611	13.00	
<i>Frequency of parent-child conversations (from 1 to 3)</i>						
Conversations about homework	23,076	2.88	0.35	22,008	2.74	0.50
Conversations about learning	22,898	2.80	0.44	21,853	2.56	0.59
Conversations about classmates	22,770	2.58	0.57	21,780	2.44	0.64
Conversations about lessons	22,766	2.58	0.56	21,813	2.45	0.61
Conversations about academic future	22,772	2.40	0.64	21,936	2.73	0.48
Conversations about professional future	22,683	2.23	0.70	21,973	2.65	0.54
<i>Cultural objects in the house (from 0 to 5)</i>						

Chapter 2 - Intelligence and academic skills

Number of books in the household	22,972	2.39	1.07	22,972	2.39	1.07
Number of CDs in the household	22,684	2.25	0.93	22,684	2.25	0.93
<i>Extracurricular activities (%)</i>						
Sport club	22,167	58.00		21,257	53.00	
Library	21,812	50.00		20,652	36.00	
Music conservatory	21,279	19.00		20,437	17.00	
Theatre class	21,076	6.00		20,206	5.00	
Scouts	21,046	4.00		20,185	4.00	

Note: Frequencies of parents-child conversations are measured on a scale from 1 to 3 (1="Never or almost never", 2="From time to time", 3="Regularly"); and numbers of books and CDs on a scale from 0 to 5 (0="None", 1="between 1 and 29", 2="between 30 and 99", 3="between 100 and 199", 4="200 or more")

Predictors of the IQ-Achievement Gap in Grades 6 and 9

Introducing non-verbal IQ in Models M'6 and M'9 increased the variance of achievement explained from 42% (**Table 3**, Model M6) to 59% (**Table 3**, Model M'6) in grade 6, and from 44% (**Table 3**, Model M9) to 63% in grade 9 (**Table 3**, Model M'9). Non-verbal IQ had the largest coefficient value of all predictors: an increase of one standard deviation in non-verbal IQ increased achievement by about half a standard deviation in grades 6 and 9 ($p < 0.001$). Besides, including non-verbal IQ reduced the coefficients of almost all predictors in absolute value. The only predictor for which we found the opposite effect was sex (from -0.116 without non-verbal IQ to -0.173 with non-verbal IQ in grade 6, and from -0.084 to -0.109 in grade 9; all $p < 0.001$). We defined the IQ-achievement gap as the difference between actual achievement and that predicted by IQ. It is positive when achievement exceeds what is predicted from IQ. According to our statistical modeling (Models M'6 and M'9), a positive coefficient attached to a given predictor indicates that this predictor contributes to having a higher achievement than predicted by IQ (controlling for other predictors). Therefore, a predictor is associated with a positive IQ-achievement gap when its coefficient in Models M'6 and M'9 is positive. Perceived self-efficacy, parental education, cultural objects in the house, household income, extracurricular activities and conversations between parents and child about school induced a significantly positive IQ-achievement gap (ordered by decreasing effect size). Thus, an increase of one standard deviation in self-efficacy and parental education both increased achievement by about 0.2 of a standard deviation in grade 6 and grade 9. Conversely, being in a priority education school, female sex, meeting with teachers, commitment to help the child, and parents-child conversations about the future induced a significantly negative IQ-achievement gap (ordered by decreasing effect size). Thus, being in a priority education

school decreased achievement by about 0.2 of a standard deviation in grade 6 and grade 9. Parents' origins and parental involvement in school life had no significant influence on the IQ-achievement gap. Estimates of the full model (including factor loadings for achievement and self-efficacy indicators) are displayed in **Table B1**. **Figure 2** illustrates to what extent individual and contextual predictors in the model explain the IQ-achievement gap in grade 6.

Predictors of Academic Progression During Middle School

Our longitudinal model explained 87% of the variance in achievement in grade 9. The large coefficient for Achievement in grade 6 (0.85, $p < 0.001$) shows that our latent construct Achievement is very stable over time (**Table 4**). As a result, the effects of other predictors are very small (lower than 0.10). Non-verbal IQ in grade 6 significantly predicted progression in achievement throughout middle school, but its effect was small ($\beta = 0.021$; $p < 0.001$). Change in achievement over middle school was mostly affected by school environment: attending a priority education school in grade 6 reduced achievement in grade 9 by 0.07 standard deviation ($p < 0.001$). Parental education had a moderate and significantly positive influence on academic progression ($\beta = 0.05$), while household income, cultural objects in the house, and parental involvement in school life had a significantly positive but small impact. Likewise, conversations about future and meetings with teachers had a significant negative influence on progression, but their effects were small. Extracurricular activities, conversations about school, and commitment to help the child had no significant effect, as well as parents' origins, self-efficacy and sex. Estimates of the full model (including regression coefficients from predictors in grade 6 to achievement in grade 6, and factor loadings for achievement and self-efficacy indicators) are displayed in **Table B2**.

Table 3*Estimates for Models of Achievement Predictors in Grades 6 and 9*

	Achievement in grade 6		Achievement in grade 9	
	Model M6 R ² = 0.42	Model M'6 R ² = 0.59	Model M9 R ² = 0.44	Model M'9 R ² = 0.63
Regression coefficients				
Non-verbal IQ		0.473* (0.005)		0.493* (0.005)
Self-efficacy	0.329* (0.008)	0.254* (0.007)	0.288* (0.008)	0.216* (0.006)
Female	-0.116* (0.012)	-0.173* (0.011)	-0.084* (0.012)	-0.109* (0.01)
Household income	0.076* (0.009)	0.061* (0.01)	0.1* (0.01)	0.076* (0.008)
Parental education	0.256* (0.008)	0.204* (0.007)	0.268* (0.008)	0.212* (0.007)
Cultural objects in the house	0.179* (0.007)	0.128* (0.006)	0.18* (0.007)	0.128* (0.006)
Priority education school	-0.289* (0.016)	-0.197* (0.015)	-0.305* (0.016)	-0.221* (0.014)
Extracurricular activities	0.047* (0.006)	0.04* (0.005)	0.077* (0.006)	0.06* (0.005)
Parents born abroad	-0.076 (0.022)	-0.026 (0.019)	-0.137* (0.022)	-0.06 (0.018)
Conversations about school	0.031* (0.007)	0.033* (0.006)	0.025* (0.007)	0.025* (0.005)
Conversations about future	-0.076* (0.006)	-0.051* (0.006)	-0.084* (0.007)	-0.057* (0.005)
Involvement in school life	0.026* (0.006)	0.012 (0.005)	0.035* (0.006)	0.021* (0.005)
Meeting with teachers	-0.145* (0.006)	-0.095* (0.005)	-0.159* (0.007)	-0.093* (0.005)
Child help	-0.125* (0.006)	-0.089* (0.005)	-0.112* (0.006)	-0.064* (0.005)

Note: Standardized estimates are reported. Standard errors are in parenthesis. *p<0.001. N = 23,258. Non-verbal IQ, Self-efficacy, Household income, Priority education school, Extracurricular activities, Conversations about school, Conversations about future, Involvement in school life, Child help, and Meeting teachers are measured in grade 6 in Models M6 and M'6, and in grade 9 in Models M9 and M'9. Model M6: RMSEA = 0.068, CFI=0.847; Model M'6: RMSEA = 0.071 CFI = 0.836; Model M9: RMSEA = 0.066, CFI = 0.865; Model M'9: RMSEA = 0.072, CFI = 0.849. Source: MENESR DEPP

Figure 2

Relation between Actual Achievement and Predicted Achievement in Grade 6

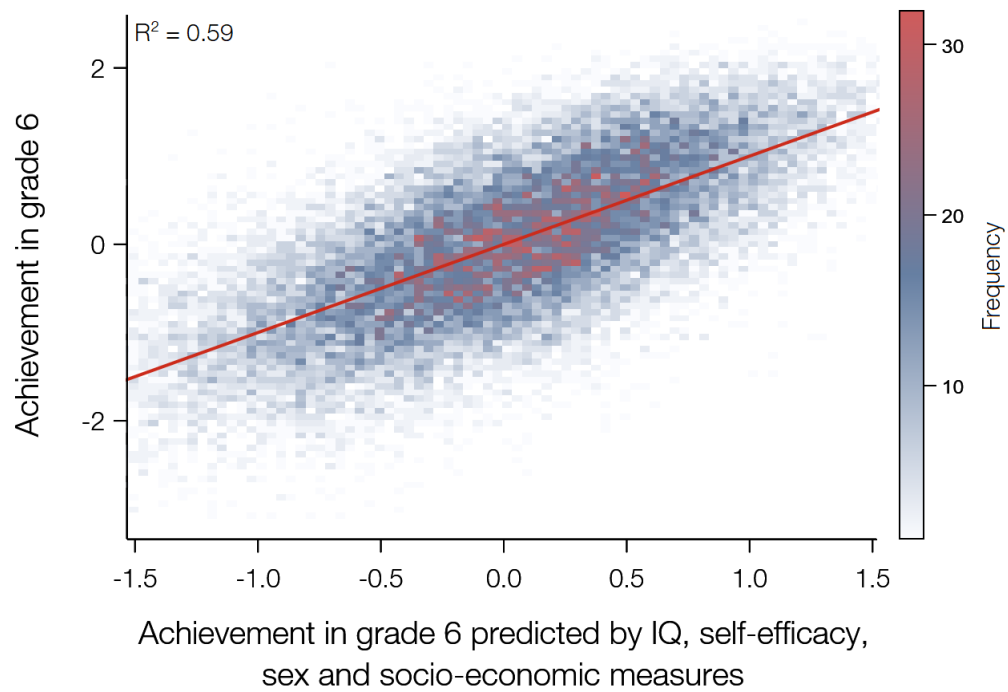
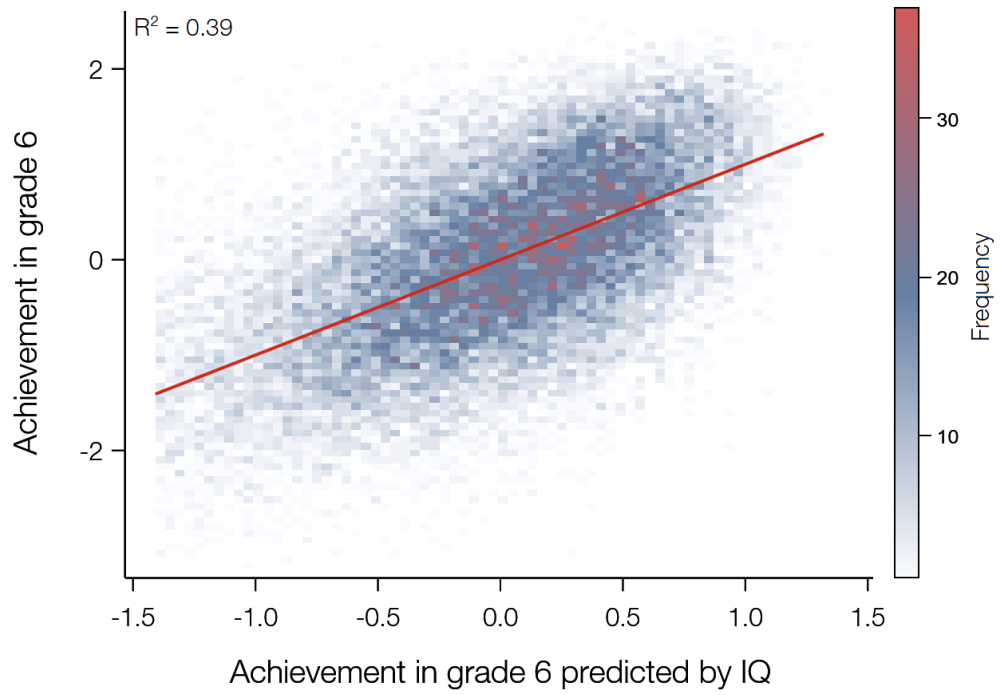


Table 4*Effects of Predictors in Grade 6 on Academic Progression Between Grade 6 and Grade 9*

	Estimates of the Longitudinal Model	
	R ² for Achievement in grade 9 = 0.87	
	β	SE
Predictors of Achievement in grade 9		
Achievement in grade 6	0.851*	0.005
Non-verbal IQ	0.021*	0.005
Self-efficacy	0.015	0.005
Female	0.01	0.007
Household income	0.029*	0.006
Parental education	0.048*	0.005
Cultural objects in the house	0.022*	0.004
Priority education school	-0.07*	0.01
Extracurricular activities	0	0.004
Parents born abroad	0.008	0.013
Conversations about school	0.016*	0.004
Conversations about future	-0.017*	0.004
Involvement in school life	0.014*	0.004
Meeting with teachers	-0.027*	0.004
Child help	-0.004	0.004

Note: Standardized estimates are reported. *p<0.001. N = 23,258. RMSEA = 0.062, CFI = 0.904.

2.1.5 Discussion

The aim of this study was to assess the contributions of various environmental and individual factors to the IQ-achievement gap – i.e., the part of achievement that is not predicted by IQ – during middle school. Our study confirmed that cognitive ability is an important predictor of academic achievement, even when taking into account a wide range of contextual and individual factors. Indeed, non-verbal IQ was by far the most predictive factor of achievement in grades 6 and 9, and not including it induced biased estimates for other independent variables. However, it had a small effect on the variations in achievement

during middle school compared to other factors. These results thus confirm the undeniable and widely reported role of intelligence (here non-verbal intelligence) in explaining academic performance at a given time, but show that it has a marginal role in explaining progression.

Environment and the IQ-Achievement Gap

School environment, parental education and cultural objects in the house were the environmental factors that had the strongest influence on academic achievement, independently of non-verbal IQ (coefficients respectively equal to -0.2, 0.2, and 0.13) (Hypothesis 1a). These factors also significantly (but modestly) affected academic progression over middle school, thus showing their long-lasting and cumulative effects (coefficients were small, equal to -0.07, 0.05 and 0.02, respectively) (Hypothesis 2). Students from highly educated families and those who had access to cultural resources at home not only succeeded more than expected given their non-verbal IQ in grades 6 and 9, but also seem to have progressed more than their peers, thereby increasing the gap. Household income also positively affected the IQ-achievement gap and academic progression, although more moderately, while extracurricular activities were positively associated with the IQ-achievement gap but did not affect progression.

Conversely, students enrolled in middle schools in disadvantaged areas in grade 6 performed lower than expected in grades 6 and 9, and regressed compared to their peers, although RAR and RRS schools received additional State funding and personnel. Our data does not allow us to determine whether these additional resources had the intended effects because the corresponding variables were not available. However, our results suggest that the additional resources allocated to the RAR and RRS schools certainly did not suffice to compensate all the difficulties linked to disadvantaged areas. These results echo those of Bénabou, Kramarz, & Prost (2009), who estimated the impact of the former *Zones d'Education Prioritaire* program (Priority Education Zones; that preceded the RAR and RRS) on academic achievement in the 1980s. They had found that in spite of decreases in class size and increases in teaching hours per student in the concerned schools, the effect was nil. Such findings point towards the need for further research on the impact of more recent priority school programs, in order to improve the allocation of resources and their effect.

Chapter 2 - Intelligence and academic skills

Contrary to our expectations, the effect of parental involvement indices was overall negative. On the one hand, conversations between parents and child about school and parental involvement in school life were positively related with the IQ-achievement gap and academic progression – which is in line with past research (Cheung & Pomerantz, 2012; Hill & Craft, 2003) –, but their effects were small or non-significant (less than 0.03). On the other hand, conversations with the child about professional future, help provided to the child, and meetings with teachers were negatively associated with the IQ-achievement gap, and negatively influenced academic progression (even though their effect on progression was small or non-significant – coefficients lower than 0.1). Thus, these variables may reflect a combination of the expected positive effect and of the opposite one, i.e., that children with more difficulties (that were not entirely captured by non-verbal IQ nor achievement, such as behavioral problems) generate more conversations about their schooling and their future, need more help with homework, and more meetings between parents and teachers. The net effect being negative suggests that the latter relationship is predominant.

These associations between family and school environments and achievement are likely to be partly mediated by genetic factors. Indeed, heritability studies on a large sample of UK twins have shown that common genetic influences accounted for 59% of the correlation between results at a national examination and perceptions of school environment (Krapohl et al., 2014), and half the correlation between achievement and family SES (Krapohl & Plomin, 2016).

Perceived Self-Efficacy and the IQ-Achievement Gap

Perceived self-efficacy proved to be one of the most predictive factors of academic achievement in grades 6 and 9 after non-verbal IQ (effect around 0.2) (Hypothesis 1b). However, it did not significantly affect academic progression during middle school, contrary to our hypothesis (Hypothesis 2). Hence, students with higher levels of perceived self-efficacy had higher results than expected from their non-verbal IQ, but did not progress more than their peers. This effect of self-efficacy on academic achievement is likely to be partly accounted for by genetic factors, as more than 60% of the association between self-efficacy and achievement is mediated genetically (Greven et al., 2009; Krapohl et al., 2014). However, in the case of cross-sectional models, it may also be noted to be partly

circular, in the sense that children's achievements, as attested by their grades, probably affected their perceived self-efficacy, since both were measured at the same time.

Besides, our cross-sectional and longitudinal models showed that non-verbal IQ in grades 6 and 9 predicted results in standardized academic tests better than did self-efficacy. This result is consistent with those of a cross-cultural (Dutch, UK and US) study conducted by Borghans, Golsteyn, Heckman, & Humphries (2016), who found that IQ predicted results at standardized achievement tests better than did personality, while personality predicted teacher grades better than did IQ.

Sex Differences

Our results indicated that girls under-performed (β lower than -0.1) compared to what could be expected from their non-verbal IQ in grades 6 and 9 (Hypothesis 1c). Interestingly, while the effect of being a girl was negative in models that did not include non-verbal IQ as a predictor (**Table 3**, Models M6 and M9), it became even more negative once non-verbal IQ was added (**Table 3**, Models M'6 and M'9). This is due to the fact that in this population, girls scored slightly higher than boys in non-verbal IQ ($d = 0.115$ in grade 6 and 0.046 in grade 9). Given this higher non-verbal IQ, higher level of achievement was expected, which was not the case. However, overall progression in academic skills during middle school was not affected by sex; thus, Hypothesis 2 was not confirmed for sex. These results contrast with those of Deary et al. (2007), who had found that English girls performed better than boys, independently from IQ. Such difference may partly come from cultural differences, from the nature of the academic tests considered, or even from the absence of environmental factors and self-efficacy indicators in their study.

Limitations

Our results may be weakened by the lack of several factors that could influence the IQ-achievement gap, such as genetic factors, parental IQ, early cognitive abilities and behavioral or emotional difficulties, whose absence may have biased some of the coefficients. Genetic factors are a first important confounder. Indeed, there exists a significant genetic overlap between perceived school environment, home environment, intelligence and achievement (Krapohl et al., 2014). Genes have also been found to mediate up to half the correlation between achievement scores and family SES, with only one third of this association accounted for by general intelligence (Krapohl et al., 2014). Similarly, a significant part of the genetic covariance between self-perceived abilities and achievement

Chapter 2 - Intelligence and academic skills

is independent of IQ (Greven et al., 2009). Therefore, the effects of parental education, school environment and self-efficacy are likely to be overestimated due to the absence of genetic factors. Likewise, the absence of behavioral problem indicators may have led us to overestimate the coefficient of non-verbal IQ, as hyperactivity and conduct problems (which are negatively associated with IQ) negatively affect achievement independently of IQ (Breslau et al., 2009b; Fergusson & Horwood, 1995).

Besides, our data did not include other factors that have an effect on achievement and that may also have an effect on intelligence independently from IQ, such as implicit theories of intelligence: whether a student holds a ‘fixed mindset’ (believes that intelligence and ability in different subjects is a fixed trait – *entity theory*) or a ‘growth mindset’ (believes that intelligence is malleable and can be developed – *incremental theory*) may significantly affect their progression in achievement. Indeed, Blackwell, Trzesniewski, & Dweck (2007) showed that having a growth mindset predicts an increase in grades during middle school, while having a fixed mindset predicts a flat trajectory (but they did not control for IQ). This result was confirmed by a classroom intervention teaching the incremental theory to students: while students from the control group experienced a continuing decline in grades, students from the experimental group experienced a positive rebound. Lastly, our parental SES and priority education school variables may have incompletely captured the effects of community, school or class, which also have an impact on achievement (Burke & Sass, 2013; McEwan, 2003; Zimmerman, 2003), possibly independently of the IQ level.

Lastly, our study was also limited by the fact that only non-verbal IQ was measured. Verbal IQ reflecting more crystallized intelligence and being therefore closer to achievement scores, the part of achievement variance explained by total IQ might have been even greater than explained here by the Chartier’s Reasoning Test on Playing Cards.

Conclusion

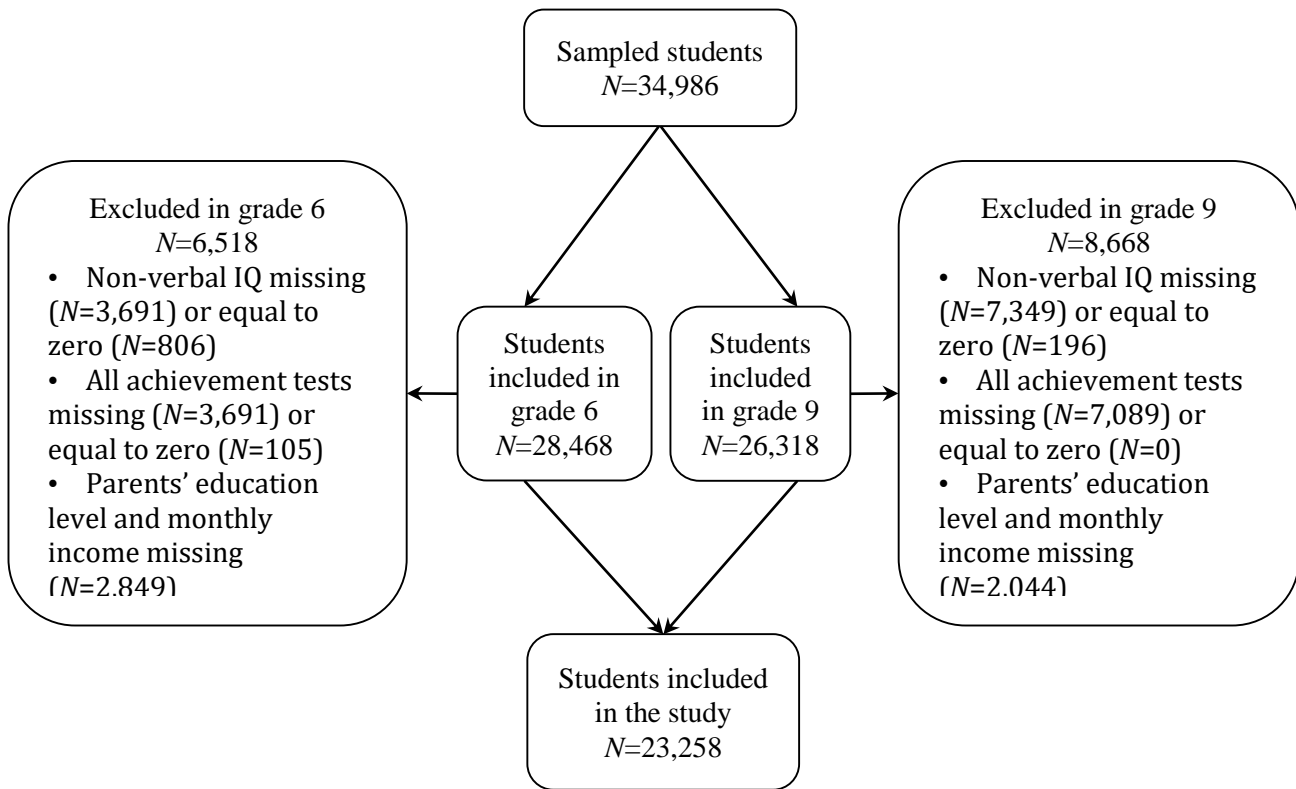
Using a large longitudinal cohort of French middle school students, our study allowed us to cast light on the relative contributions of a broad range of environmental and individual factors explaining why students under- or over-perform relative to what can be expected from their non-verbal IQ. Results showed that perceived self-efficacy, school environment, parental education and sex were the factors that most affected academic achievement independently of non-verbal intelligence cross-sectionally – although

intelligence was by far the best predictor. Longitudinally, school context and parental education best predicted progression in achievement between grade 6 and grade 9, while intelligence and other factors played a minor role. Our results thus demonstrate the powerful role of the socio-economic context, which not only has a major influence on academic achievement beyond intelligence, but also affects academic progression during middle school, such that inequalities keep growing.

2.1.6 Appendix A: Data

Figure A1

Flowchart of Included and Excluded Participants



Note: Source: MENESR DEPP

Table A1*Internal Consistency for Scores in Grade 6 and Grade 9*

Test	Cronbach's alpha	
	Grade 6	Grade 9
Non-verbal IQ	0.876	0.869
Academic knowledge	0.856	0.884
Reading comprehension	0.743	0.816
Mathematics	0.906	0.935
Grammar	0.831	0.817
Phonics skills	0.672	0.641
Academic self-efficacy	0.863	0.864
Social self-efficacy	0.805	0.801
Self-regulation	0.894	0.809

Table A2*Results from Exploratory Factor Analyses (EFA) with Achievement Scores in Grades 6 and 9*

Test	Factor loadings in	Factor loadings in
	grade 6	grade 9
Phonics skills	0.59	0.59
Grammar	0.83	0.85
Mathematics	0.79	0.79
Reading comprehension	0.73	0.78
Academic knowledge	0.74	0.82

Table A3*Results from Exploratory Factor Analyses (EFA) with Self-Efficacy Scores in Grades 6 and 9*

Scores	Factor loadings in	Factor loadings in
	grade 6	grade 9
Perceived self-regulation	0.42	0.47
Perceived academic self-efficacy	0.75	0.81
Perceived social self-efficacy	0.69	0.41

Table A4

Factor Structure from the Principal Component Analysis With Promax Rotation of Parental Involvement Variables in Grade 6

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Parents attended a parent-teacher meeting	0.098	0.021	0.143	-0.332	0.633
Parents met teacher on their own initiative	0.106	0.084	0.075	0.755	0.049
Parents met teacher on the initiative of the teacher	-0.064	0.063	-0.095	0.743	-0.043
Parents part of a parents' association	0.089	-0.016	0.897	-0.025	0.055
Parents are class representatives	0.093	-0.003	0.897	-0.028	0.046
Private tutoring	-0.052	0.128	0.001	0.369	0.424
Help provided to student at home	0.174	-0.031	-0.026	0.115	0.662
Conversations about homework	0.746	0.141	0.024	0.035	0.221
Conversations about learning	0.823	0.234	0.062	0.023	0.179
Conversations about classmates	0.750	0.331	0.137	-0.045	-0.009
Conversations about lessons	0.738	0.432	0.121	-0.013	0.032
Conversations about academic future	0.364	0.919	-0.005	0.113	0.028
Conversations about professional future	0.291	0.927	-0.013	0.078	0.006

2.1.7 Appendix B : Results

Table B1*Estimates for Models of the Achievement Predictors in Grades 6 and 9*

	Achievement in grade 6		Achievement in grade 9	
	Model M6 R ² = 0.42	Model M'6 R ² = 0.59	Model M9 R ² = 0.44	Model M'9 R ² = 0.63
Regression coefficients				
Non-verbal IQ		0.473* (0.005)		0.493* (0.005)
Self-efficacy	0.329* (0.008)	0.254* (0.007)	0.288* (0.008)	0.216* (0.006)
Female	-0.116* (0.012)	-0.173* (0.011)	-0.084* (0.012)	-0.109* (0.01)
Household income	0.076* (0.009)	0.061* (0.01)	0.1* (0.01)	0.076* (0.008)
Parental education	0.256* (0.008)	0.204* (0.007)	0.268* (0.008)	0.212* (0.007)
Cultural objects in the house	0.179* (0.007)	0.128* (0.006)	0.18* (0.007)	0.128* (0.006)
Priority education school	-0.289* (0.016)	-0.197* (0.015)	-0.305* (0.016)	-0.221* (0.014)
Extracurricular activities	0.047* (0.006)	0.04* (0.005)	0.077* (0.006)	0.06* (0.005)
Parents born abroad	-0.076 (0.022)	-0.026 (0.019)	-0.137* (0.022)	-0.06 (0.018)
Conversations about school	0.031* (0.007)	0.033* (0.006)	0.025* (0.007)	0.025* (0.005)
Conversations about future	-0.076* (0.006)	-0.051* (0.006)	-0.084* (0.007)	-0.057* (0.005)
Involvement in school life	0.026* (0.006)	0.012 (0.005)	0.035* (0.006)	0.021* (0.005)
Meeting with teachers	-0.145* (0.006)	-0.095* (0.005)	-0.159* (0.007)	-0.093* (0.005)
Child help	-0.125* (0.006)	-0.089* (0.005)	-0.112* (0.006)	-0.064* (0.005)
Factor loadings for Achievement				
Phonics skills	0.573* (0.005)	0.569* (0.005)	0.567* (0.005)	0.566* (0.005)
Grammar	0.815* (0.003)	0.795* (0.003)	0.83* (0.003)	0.816* (0.003)
Mathematics	0.79* (0.003)	0.822* (0.003)	0.791* (0.003)	0.821* (0.003)
Reading comprehension	0.711* (0.004)	0.7* (0.004)	0.762* (0.003)	0.752* (0.003)
Academic knowledge	0.734* (0.004)	0.72* (0.004)	0.818* (0.003)	0.803* (0.003)
Factor loadings for Self-efficacy				
Autoregulation	0.457* (0.008)	0.446* (0.008)	0.48* (0.009)	0.467* (0.009)
Academic self-efficacy	0.793* (0.007)	0.793* (0.008)	0.816* (0.011)	0.833* (0.012)
Social self-efficacy	0.628* (0.007)	0.636* (0.007)	0.393* (0.007)	0.392* (0.007)

Note: Standardized estimates are reported. Standard errors are in parenthesis. * $p < 0.001$. $N = 23,258$.

Non-verbal IQ, Self-efficacy, Household income, Priority education school, Extracurricular activities, Conversations about school, Conversations about future, Involvement in school life, Child help, and Meeting teachers are measured in grade 6 in Models M6 and M'6, and in grade 9 in Models M9 and M'9. Model M6: RMSEA = 0.068, CFI=0.847; Model M'6: RMSEA = 0.071 CFI = 0.836; Model M9: RMSEA = 0.066, CFI = 0.865; Model M'9: RMSEA = 0.072, CFI = 0.849. Source: MENESR DEPP

Table B2*Effects of Predictors in Grade 6 on Academic Progression Between Grade 6 and Grade 9*

	Estimates of the Longitudinal Model	
	R ² for Achievement in grade 9 = 0.87	
	β	SE
Predictors of Achievement in grade 9		
Achievement in grade 6	0.851*	0.005
Non-verbal IQ	0.021*	0.005
Self-efficacy	0.015	0.005
Female	0.01	0.007
Household income	0.029*	0.006
Parental education	0.048*	0.005
Cultural objects in the house	0.022*	0.004
Priority education school	-0.07*	0.01
Extracurricular activities	0	0.004
Parents born abroad	0.008	0.013
Conversations about school	0.016*	0.004
Conversations about future	-0.017*	0.004
Involvement in school life	0.014*	0.004
Meeting with teachers	-0.027*	0.004
Child help	-0.004	0.004
Predictors of Achievement in grade 6		
Non-verbal IQ	0.455*	0.005
Self-efficacy	0.256*	0.007
Female	-0.109*	0.011
Household income	0.051*	0.008
Parental education	0.207*	0.007
Cultural objects in the house	0.134*	0.006
Priority education school	-0.196*	0.015
Extracurricular activities	0.041*	0.005
Parents born abroad	-0.036	0.02
Conversations about school	0.037*	0.006
Conversations about future	-0.049*	0.006
Involvement in school life	0.01	0.005
Meeting with teachers	-0.093*	0.005

Chapter 2 - Intelligence and academic skills

Child help	-0.088*	0.005
------------	---------	-------

Factor loadings for Achievement in grade 9

Phonics skills	0.578*	0.005
Grammar	0.833*	0.002
Mathematics	0.792*	0.003
Reading comprehension	0.763*	0.003
Academic knowledge	0.801*	0.003

Factor loadings for Achievement in grade 6

Phonics skills	0.564*	0.005
Grammar	0.81*	0.003
Mathematics	0.811*	0.003
Reading comprehension	0.705*	0.004
Academic knowledge	0.727*	0.003

Factor loadings for Self-efficacy in grade 6

Autoregulation	0.447*	0.008
Academic self-efficacy	0.796*	0.008
Social self-efficacy	0.63*	0.007

Correlations between the same academic skills in grade 6 and grade 9

Phonics skills	0.303*	0.006
Grammar	0.28*	0.008
Mathematics	0.64*	0.005
Reading comprehension	0.11*	0.007
Academic knowledge	0.469*	0.006

Note: Standardized estimates are reported. * $p < 0.001$. $N = 23,258$. $RMSEA = 0.062$, $CFI = 0.904$.

Source : MENESR DEPP

2.2 Are high-IQ students more at risk of school failure?

This section is based on the following published article:

Guez, A., Peyre, H., Le Cam, M., Gauvrit, N., & Ramus, F. (2018). Are high-IQ students more at risk of school failure? *Intelligence*, 71, 32–40.

2.2.1 Abstract

While it is well-established that intelligence tests positively predict academic achievement, there remain widespread beliefs that gifted students experience difficulties at school and are particularly at risk of school failure. Many studies have provided evidence to the contrary, however few were based on representative population samples. This paper intended to assess whether prior results on the academic success of gifted children could be generalized to a large sample from the general French population. We analyzed a database of French middle school students (N=30,489), including scores in a fluid intelligence test in grade 6 and a variety of school performance measures in grade 9 (results at a national exam, teachers' grades, academic orientation in high school). In addition, self-efficacy and motivation were assessed. Our results replicate and extend previous findings: high-IQ students scored much better on all academic performance measures, which was corroborated by higher levels of motivation and self-efficacy. Consistently with the previous literature, there was a robust positive relationship between fluid intelligence in grade 6 and academic performance in grade 9 in the whole sample, which was also observed within high-IQ students. Exploratory analyses revealed that IQ moderated the association between social background and children's achievement, such that the positive link between parental education and achievement levelled off for high-IQ children. The positive association between high-IQ and achievement was similar for boys and girls.

Keywords: Giftedness, high-IQ, academic achievement, intelligence

2.2.2 Introduction

Intelligence tests were originally designed with the explicit purpose of predicting school success (Binet & Simon, 1904). Since then, after a century of further development of tests and theorising, scores provided by intelligence tests remain a robust predictor of academic achievement (Deary et al., 2007; Rohde & Thompson, 2007; Roth et al., 2015). More generally, IQ is positively correlated with a large array of life outcomes, including income (Zagorsky, 2007), mental and physical health (Der et al., 2009; Gale et al., 2009), or life expectancy (Batty et al., 2007).

In this context, it may seem surprising that there remain widespread beliefs about gifted children suffering from social and emotional difficulties. For example, the National Association for Gifted Children states that gifted children “may be at greater risk for specific kinds of social-emotional difficulties if their needs are not met”, such as “heightened awareness, anxiety, perfectionism, stress, issues with peer relationships, and concerns with identity and fit” (*Social & Emotional Issues / National Association for Gifted Children*, s. d.). Similarly, on the website of the National Register of Health Service Psychologists, James T. Webb writes that many professionals “are unaware that talented and gifted children are at risk for underachievement, peer relationship issues, power struggles, perfectionism, existential depression, and other problems, and that bright adults often have job difficulties, problems with peers, spouses or children, and existential depression that stem from giftedness.” (Webb, 2014). These beliefs are supported by studies that show positive associations between high IQ and anxiety (Lancon et al., 2015), depression (Jackson & Peterson, 2003), internalizing and externalizing problems (Guérolé et al., 2013) and various psychological and physiological disorders (Karpinski et al., 2018). However, these studies relied on case studies or biased samples (such as members of Mensa in Karpinski et al. (2018)⁸, or clinically referred children in Guérolé et al.(2013)).

Even more surprising, some people seem to think that gifted children are more at risk of school failure, potentially due to the above mentioned social and emotional problems, but also to lower self-efficacy or motivation (Reis & McCoach, 2000), heightened risk of bullying (Peterson & Ray, 2006), boredom in class (Vannetzel, 2009) or perfectionism (Webb, 2014) – even though some studies argue against these hypotheses

⁸ For detailed comments on this article see <https://pubpeer.com/publications/2F26A22D54A2032B460B3037AF26C0>

(Feldhusen & Kroll, 1991; McCoach & Siegle, 2003; Peters & Bain, 2011; Roznowski et al., 2000). Other authors argue in favour of the existence of a “negative Pygmalion effect”, that would encourage the child to conform to its environment and the lower demands of the school in order to be accepted by others, which, as a consequence, would increase socio-emotional problems and heighten the risk of failure (Terrassier, 2009). Thus, popular media report that 20% of gifted student may drop out of school in the US (Kuzujanakis, 2013), while in France, the reported proportion of gifted children failing at school goes from one third (Bourgeois, 2017; Colonat, 2018; Le Saint, 2017) to up to 70% (Quillet, 2012). Here again, these figures are supported by little evidence, or come from biased samples – e.g. the estimate of one third of failing gifted students in France comes from a survey of parents of children belonging to the French Association for Gifted Children (Côte, 2005).

In contrast, scientific evidence converges towards the fact that gifted students perform better than their peers. The literature on the achievement of gifted students goes back to the 1920s with Terman’s Study of the Gifted (Terman, 1926a). This longitudinal study examined the characteristics and development of 1,528 high-IQ children in California, aged 2 to 13 at the beginning of the study. Gifted students from the main experimental group were selected in Californian public schools by the means of a three-step process involving teacher nomination, the National Intelligence Test, and an abbreviated version of the Stanford-Binet test (belonging to the top 1%). Results showed that the gifted participants were rated higher by teachers on the quality of their school work compared to a control group (Terman, 1926b), and performed better at the Stanford Achievement Tests by two to five times the standard deviation of the controls (Terman, 1926c). Starting in the 1970s, the Study of Mathematically Precocious Youth (SMPY) followed five cohorts of American gifted students. The first three cohorts were identified at 12-13 years old by talent searches and selected with scores at the mathematics and verbal subtests of the SAT (Lubinski & Benbow, 2006). The first cohort included 2,188 students in the top 1%, the second, those in the top 0.5% (N=778), the third, those in the top 0.01% (N=501). By age 33, 25% of participants of the first cohort had earned a doctorate, 30% of cohort 2, and 50% of cohort 3 – compared to 1% in the general population (Lubinski & Benbow, 2006). Similarly, McCoach & Siegle (2003) have shown that gifted university students in the US (identified by school district volunteers) have higher self-reported Grade Point Averages (GPAs) than students from the general population, but their sample was small and not representative. Matthews (2006) also reported that in North Carolina, less than 1% of gifted high-school students (as identified by a talent search) dropped out.

However, a common limitation to these studies is their respective selection process. Indeed, they relied on teacher nomination or talent searches, which may have favoured the inclusion of academically successful gifted students in the gifted sample at the expense of low achievers, thus potentially amplifying the difference between gifted and controls. Overcoming this limitation, Roznowski et al. (2000) led a large scale study examining various academic outcomes among 12,630 American gifted and non-gifted students from the general population. Their results show that gifted individuals (top 5%) are more likely to participate in college preparatory programs, receive A and B grades in school, spend more time on homework, be less absent, like school more, feel more at ease in academic courses, and have higher self-esteem. However, their measure of cognitive ability relied on highly academic skills (they used a composite score of vocabulary, reading and arithmetic tests – BYTEST), which considerably reduces the strength of their results.

While the belief that high-IQ student are more at risk of school failure has not been supported by the literature and seems contradictory with the generally positive correlation between IQ and achievement, it is not inconceivable that this relationship might reverse or at least level off beyond a certain IQ level, such that individuals with very high IQ might succeed less well than expected from the linear relationship observed in non-gifted children. Again, this threshold hypothesis has not been supported by the literature, as a series of studies showed the existence of differences in degrees earned and other indicators of success depending on ability levels, even within the highly gifted SMPY population (Park et al., 2008; Robertson et al., 2010).

These findings together indicate that gifted students, far from being worse off at school, outperform their peers. However, most of these studies involved nomination or talent search as a selection step in order to find gifted participants, which implies that the participants in these studies may be biased in favour of successful gifted children. Therefore, these results were often based on non-representative samples, so that there remains a need for research on the gifted in the general population. Besides, these studies were all conducted in the United States, which raises the question of the generalisability of their findings in other countries.

In order to test whether previous results on the academic performance of gifted students could be replicated in a large representative sample and in a different population, we analysed data from 30,489 French middle school students. Giftedness is a very broad term, which can refer to superior abilities in multiple domains, such as general intellectual ability, leadership skills, or visual/performing arts. In this paper, we investigate intellectual

giftedness, i.e. superior general intellectual ability. The data used includes scores in a fluid intelligence test in grade 6 and a variety of school performance measures in grade 9 (results at a national exam, teachers' grades, academic orientation in high school). In addition, self-efficacy and motivation were assessed. This rich database thus allowed us to study the differences in a large range of school performance measures between gifted and non-gifted students in France.

In accordance with the existing literature, we formulated the following hypotheses:

- a) High-IQ students show better academic achievement than other students.
- b) They drop out less frequently from middle school.
- c) High-IQ students show higher scores in measures of self-efficacy and motivation.
- d) There is a positive relationship between IQ in 6th grade and achievement in 9th grade.
- e) This relationship holds equally in high-IQ students and in the general population.

2.2.3 Method

Sample

We analyzed data from the DEPP Panel 2007, a study directed by the Direction de l'Évaluation, de la Prospective et de la Performance (DEPP), French Ministry of Education (Trosseille et al., 2013). 34,986 children were followed from their entrance in the first year of French middle school (grade 6) in 2007 to the second year of high school (grade 11). The study was compulsory and approved by the National Council for Statistical Information (CNIS) (visa n°2008A061ED and 2011A082ED), ensuring public interest and conformity with ethical, statistical and confidentiality standards. The sampling strategy was balanced sampling, i.e. the random selection of a sample that is representative of the sampling frame based on available characteristics, using the algorithm CUBE created by INSEE (Rousseau & Tardieu, 2004). The sample was randomly selected from an exhaustive sampling frame, the *Système d'information du second degré de la DEPP*, balancing the following characteristics: region, public/private status of the school, urban unit, school establishment, age of entry in grade 6 (Trosseille et al., 2013). The sample was constituted in such a way as to be representative of the French population of middle school students, with a slight over-representation of students in schools belonging to the Réseau Ambition Réussite (Success Ambition Network – schools in disadvantaged areas). Previous

analyses of this dataset have been conducted before (Guez et al., 2018; Trosseille et al., 2013). Our working sample includes students for whom the intelligence score in grade 6 was available and different from zero ($N = 30,489$). **Table 1** describes the main socio-demographic characteristics for high-IQ and non-high-IQ students in our sample.

Table 1*Descriptive Statistics for Sociodemographic Variables and Fluid Intelligence*

Variable	High-IQ N=888		Others N=29601		Difference (High-IQ - Others)	
	<i>M or</i>	<i>SD</i>	<i>M or %</i>	<i>SD</i>	<i>Cohen's d or</i>	<i>p</i>
	<i>%</i>				<i>Odds Ratio</i>	
RCC score (fluid intelligence) in grade 6 (out of 30)	26.75	0.92	15.06	5.83	2.03	<.0001
Age in grade 6 (years)	10.93	0.35	11.16	0.48	0.48	<.0001
Parental education (years)	13.61	3.24	11.87	3.47	0.50	<.0001
Disadvantaged school (%)	10.09		18.43		0.50	<.0001
Both parents born abroad (%)	1.87		10.53		0.58	<.0001
Sex (% Female)	51.01		49.53		1.06	.7603

Note: RCC=Chartier's Reasoning Test on Playing Cards. Cohen's d are indicated for differences in continuous variables (RCC scores, age in grade 6, years of parental education), and odds ratios for differences in proportions (%).

Measures

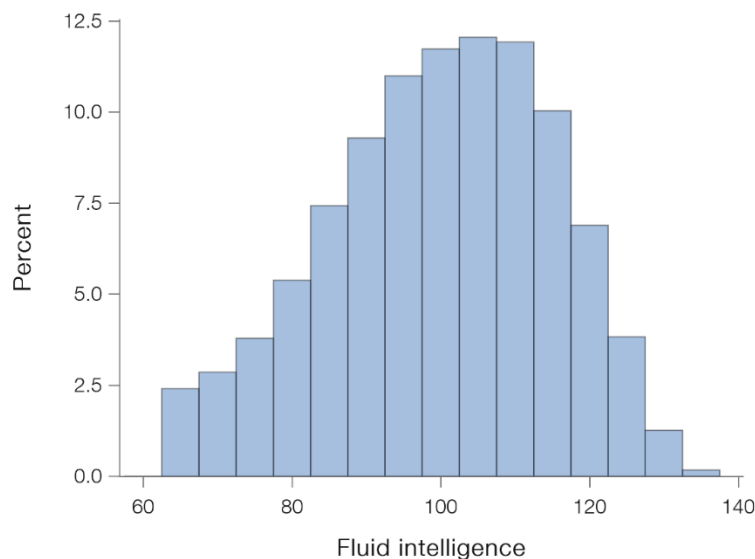
Fluid intelligence

In grade 6 (2007) and grade 9, the *Raisonnement sur Cartes de Chartier* (RCC – Chartier's Reasoning Test on Playing Cards) test was administered collectively to all participants. Designed to capture fluid intelligence (*gf*) (Chartier, 2012; Terriot, 2014), the RCC contains 30 items that evaluate children's non-verbal logical reasoning skills, inspired from Raven's progressive matrices (Raven, 1998) but using playing cards as material. Each item is solved by determining which card would fill the blank in an array composed of 4 to 12 cards. The RCC is scored as the number of items correctly completed in a limited time (20 minutes). Internal consistency was good ($\alpha = 0.88$), and the correlation between RCC scores in grade 6 and in grade 9 was relatively strong ($r = 0.61$), indicating a good reliability (similar to findings by Watkins & Smith (2013), who found correlation coefficients for WISC-IV subtests administered 3 years apart ranging from .46 to .70, with a median of

.56). A large number of students scored 0 (2.58%), many more than those scoring 1 (1.16%), which may mean that they were not engaged in the task, refused to take the test, or which may indicate problems with administration and scoring. Therefore, we chose to remove them from our analyses (see **Figure A1**). We scaled RCC scores on the standard IQ scale ($M = 100$, $SD = 15$). The distribution of scores was negatively skewed (**Figure 1**), and this remained the case after using sampling weights to make the population representative (the skewness was about equal to -0.29 in both cases). Thus the negative skew was not due to the overrepresentation of schools from disadvantaged areas. As a result, only 0.55% of students had a non-verbal IQ score higher than 130 (about 2% of examinees in a normal distribution), which is a frequent threshold above which an individual is considered gifted (Carman, 2013; Newman, 2008; the top 2% criterion is also the one applied by Mensa). Therefore, in order to have about the same proportions in our sample, we categorized as high-IQ students those with an RCC score above the 98th percentile (here equivalent to a non-verbal IQ of 126.2) ($N = 888$).

Figure 1

Distribution of Fluid Intelligence in Grade 6



Academic achievement

At the end of middle school (grade 9), all French students take a national examination, the *Diplôme national du Brevet* (hereafter referred to as DNB). The final grade at the DNB is composed of continuous monitoring throughout grade 9 (teachers'

grades in all subjects) and of results at the national examination. Raw final grades undergo a “harmonization” process, whereby grades are somewhat standardized across academic regions, and grades just below the pass threshold are raised just at the threshold. At the time of the study, the DNB examination included three anonymously graded tests: one in French, one in Mathematics and one in History and Geography. Our dataset includes results at the national examination in the three subjects, grades in the continuous monitoring, and the final DNB grade.

Academic orientation

In addition to these results, we evaluated later academic success using orientation decisions at the end of middle school and at the end of the first year of high school. In France, middle school is general and common to all students. Academic paths split at the end of grade 9, when students can be oriented, depending on their wishes, on their academic record and on the school staff decision, to general and technological high school (*Lycée général et technologique*), to vocational high school (*Lycée professionnel*), or to the preparation of a vocational diploma (*Certificat d'aptitude professionnelle* or *Brevet d'études professionnelles*). Similarly, at the end of grade 10, students in general and technological high school can either be oriented to the one of the three general tracks (*Premières générales: Première S – scientific, Première ES – economics and social science, and Première L – literature track*) or to one of the nine technological tracks, ranging from management to catering or industrial sciences and technologies (*Premières technologiques: STI2D, STD2A, STG, STL, ST2S, BT, Hôtellerie, TMD, and STAV*).

Perceived self-efficacy

In grade 9, students answered questions from the Children's Perceived Self-Efficacy scales (Bandura, 1990), closely translated into French (Blanchard et al., 2013). It is a 37-item questionnaire which measure three types of self-efficacy: perceived academic self-efficacy (19 items), i.e. students' perceived ability to manage their learning, to master different academic subjects (mathematics, science, etc...), and to fulfill parents' and teachers' expectations; perceived social self-efficacy (13 items), i.e. efficacy regarding leisure group activities, the ability to form and maintain social relationships and manage interpersonal conflicts, and self-assertiveness; and perceived self-regulatory efficacy (5 items), i.e. students' perceived ability to resist peer pressure to engage in high-risk activities (alcohol, drugs, transgressive behaviors). For each item, students had to evaluate their

ability to perform each activity using a 5-point Likert scale. Principal component analysis confirmed the three main factors structure, based on the scree plot, as in Pastorelli et al. (2001). Three perceived self-efficacy scores (academic, social, and self-regulatory) were then constructed using confirmatory factor analysis. Internal consistency was good for the three indicators (Cronbach's α above 0.80).

Motivation

Lastly, students' academic motivation was assessed in grade 9 with questions derived from the Academic Self-Regulation Questionnaire (SRQ-A) (Ryan & Connell, 1989), adapted and translated into French (Blanchard et al., 2013; Leroy et al., 2013). This is a self-report measurement which assesses individual differences in motivational styles (intrinsic motivation, identified regulation, introjected regulation, external regulation and amotivation). In this framework, intrinsic motivation is defined as the most self-determined form of motivation, where a student engages in a behavior spontaneously, out of interest and enjoyment. Identified regulation, introjected regulation and external regulation are different types of extrinsic motivation, i.e. motivational styles in which a behavior is driven by factors external to the activity itself. Lastly, amotivation refers to the absence of motivation: students perceive no link between a behavior and its outcomes. 12 items from the SRQ-A were present in the grade 9 questionnaire. The items asked students the reasons why they do their homework, work on their classwork and try to do well in school. Each item provides a possible reason that represents a certain motivational style (for example: "I do my classwork because I want to learn new things", or "I do my classwork because I'd be ashamed of myself if it didn't get done"). Each item was answered using a 5-point Likert scale. A principal component analysis on the 12 motivation items in grade 9 yielded a three factors solution (the first three eigenvalues were higher than 1, but not the fourth⁹): the first factor corresponding to intrinsic motivation, the second to amotivation, and the third to extrinsic motivation. We thus created three motivation scores (amotivation, intrinsic motivation, extrinsic motivation) using a confirmatory factor analysis. Internal consistency was average to good for the three indicators (Cronbach's α equal to 0.85 for intrinsic motivation, and 0.68 for extrinsic motivation and amotivation).

⁹ We used both the eigenvalue and the scree plot to decide the number of factors retained: the eigenvalue criterion led to 3 factors (95% confidence intervals for the third and fourth eigenvalue: [1.12;1.15] and [0.67;0.70]) and the scree plot indicated a solution with 4 factors. The 3 factor solution afforded a more coherent interpretation, thus we chose to retain it.

Analyses

Descriptive statistics

Our working sample includes students for whom the RCC score was available in grade 6 and was different from zero ($N = 30,489$). We tested the difference in the main outcome variables (academic results, orientation decision in high-school, perceived self-efficacy and motivation) between high-IQ and non-high-IQ students using a t-test (for continuous outcomes), a Chi-square test (for categorical outcomes), or Fisher's exact test (for categorical outcomes with less than 5 counts in a given cell), as appropriate. For continuous variables, we estimated the effect size of the difference using Cohen's d , which indicates the standardized difference between two means (Cohen, 1988), and we used odds ratio for categorical variables. We also compared the proportion of missing DNB grades between high-IQ and non-high-IQ students (with a Chi-square test). For all group comparisons, we assessed significance with an alpha threshold of $p < 0.0025$ to take multiple testing into account (Bonferroni corrected threshold for 20 tests).

Regression analyses

We first ran univariate regressions of grade at the DNB examination (in grade 9) on non-verbal IQ score (in grade 6). We ran this model in our entire working sample first and then in our sample of high-IQ students, in order to check whether the positive relationship between non-verbal IQ and achievement that has been widely described in the literature holds for high-IQ students. As a sensitivity analysis, we recalculated this regression at different thresholds for our high-IQ category (95% and 99%).

Further, as additional exploratory analyses, we investigated whether IQ moderated the association between social background and academic achievement, and between sex and academic achievement. It has been argued that gifted children from disadvantaged social and cultural backgrounds could indeed suffer much more from a "negative Pygmalion effect" or lack of family support, and thus be more at risk of underachievement (Terrassier, 2009). In such a case, it may be that high-IQ students' achievement is more related to their social background than that of other students, i.e. the interaction term between high-IQ and social background would be positive. This is one possible rendering of the more general hypothesis that high intellectual potential requires good environmental conditions to fully flourish, and therefore that the difference between high IQ students and others is smaller in a low SES background. Conversely, one might hypothesize that high IQ buffers against the effects of social disadvantage, which would predict a larger

difference between both groups at low SES, and therefore a negative interaction term. Similarly, we explored whether the relation between high-IQ and achievement is the same in both sexes, as some have argued that gifted boys may be more vulnerable than gifted girls (Terrassier, 2009). In their sample, Roznowski et al. (2000) had indeed found that the difference in A and B grades reported between girls and boys (in favor of girls in all ability groups) was larger among gifted students than among high-ability students. We investigated these questions by regressing results at the DNB examination on a dummy variable equal to 1 if the student is in the high-IQ group (*High-IQ*), on parental education (*SES*) and on sex (*Girl*, dummy variable equal to 1 if the student is a girl), and on two interaction variables: one between high-IQ and sex (*High-IQ* × *Sex*), and one between high-IQ and parental education (*High-IQ* × *SES*). We assessed significance with an alpha threshold of $p < 0.01$ (Bonferroni corrected threshold for 5 tests).

2.2.4 Results

Differences between high-IQ and non-high-IQ students

We first checked whether gifted students showed better academic achievement than other students (Hypothesis a). Academic achievement was assessed with results at the DNB and academic orientation after middle school. **Tables 2 and 3** present descriptive statistics for outcome variables. High-IQ students had on average statistically significantly higher scores than others in all DNB subjects (their final grade was higher by 2.6 points out of 20; $p < .0001$). Besides, only 1.66% had an average grade strictly lower than the pass threshold (10 out of 20) – compared to 15.55% among non-high-IQ students (see **Table 2** and **Figure 2**).

Such higher performance at the DNB exam logically resulted in a higher proportion of high-IQ students continuing to general and technological high school compared to others (89.49% versus 61.76%; $p < .0001$), and a lower proportion being retained in grade 9 (1.56% versus 3.05%, but not statistically significant – $p = .0165$). Guidance decisions at the end of grade 10 indicate that high-IQ students kept on doing well in high school, as 82.91% of students took a general track, compared to 63.38% of students with a non-verbal IQ lower than the 98th percentile ($p < .0001$) (see **Table 3**). Moreover, it is highly likely that this reported proportion for non-high-IQ students is over-estimated, since the share of missing data for non-high-IQ students is much higher than for high-IQ students for this particular

variable (48.74% of missing data for non-high-IQ students versus 20.27% for high-IQ students).

Our second hypothesis was that high-IQ students drop out less frequently from middle school (Hypothesis b). If high-IQ students were at high risk of school failure, it might be that they actually have dropped out from middle school even before reaching grade 9. In that case, DNB results would give positively biased information about high-IQ students' school success. However, only 5.18% of DNB results were missing for high-IQ students compared to 13.78% for others, which is inconsistent with this hypothesis.

Lastly, we had hypothesized that high-IQ students also fare better in terms of perceived self-efficacy and motivation (Hypothesis c). Results show that they indeed had statistically significantly higher levels of perceived academic self-efficacy compared to other students ($d = .36; p < .0001$), as well as perceived self-regulation ($d = .20; p < .0001$) (see **Table 2** and **Figure 3**). However, their level of social self-efficacy was slightly but not statistically significantly lower than others ($d = -.08; p = .0317$). Lastly, they reported statistically significantly higher levels of both intrinsic motivation ($d = .25; p < .0001$) and extrinsic motivation ($d = .19; p < .0001$), as well as lower levels of amotivation ($d = -.18; p < .0001$) (see **Table 2** and **Figure 4**).

Table 2

Descriptive Statistics for Grades, Self-Efficacy and Motivation Scores in Grade 9

Outcome variable	High-IQ N=888		Others N=29601		Difference High-IQ - Others		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>Cohen's d</i>	99.75% C.I.	<i>p</i>
<i>DNB results (out of 20)</i>							
DNB Examinations	13.13	2.73	10	3.19	0.99	[.88 ; 1.06]	<.0001
DNB Teacher grades	15.11	2.2	12.79	2.72	0.86	[.85 ; .93]	<.0001
DNB Final grade	14.48	2.29	11.88	2.88	0.91	[.82 ; .98]	<.0001
<i>Perceived self-efficacy (factor scores)</i>							
Academic	-0.19	0.89	-0.57	1.05	0.36	[.27 ; .46]	<.0001
Social	-0.06	0.9	0.02	0.95	-0.08	[-.18 ; .03]	0.0408
Self-regulatory	0.31	0.51	0.17	0.7	0.20	[.12 ; .29]	<.0001
<i>Academic motivation (factor scores)</i>							
Intrinsic motivation	0.22	0.86	-0.01	0.94	0.25	[.14 ; .35]	<.0001
Extrinsic motivation	0.12	0.8	0	0.86	0.14	[.04 ; .25]	<.0001
Amotivation	-0.16	0.76	0	0.9	-0.18	[-.27 ; -.08]	<.0001

Note : Bonferroni corrected threshold for significance is 0.0025.

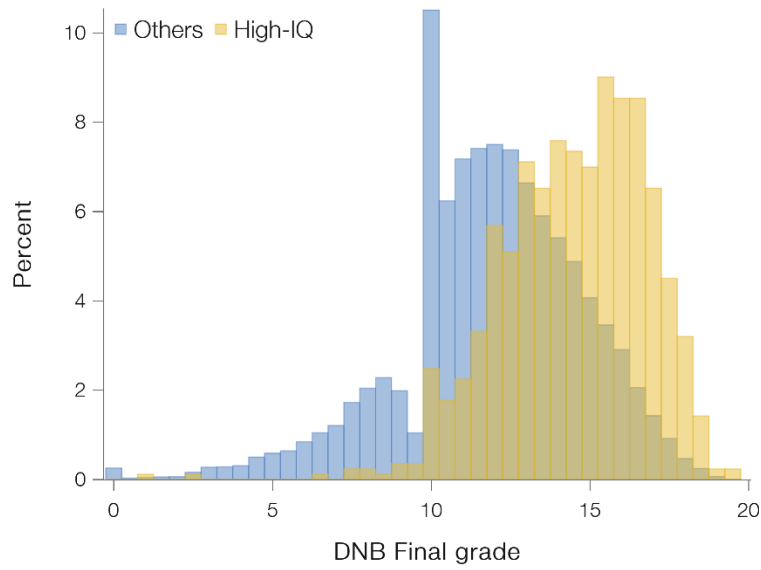
Table 3*Descriptive Statistics for Academic Orientation after Middle School*

Outcome variable	High-IQ		Others		Difference (High-IQ – Others)		
	<i>N</i>	%	<i>N</i>	%	<i>Odds Ratio</i>	99.75% C.I.	<i>p</i>
<i>Guidance decision at the end of grade 9</i>	771		25278				
General & technological high school		89.49		61.76	5.28	[3.69 ; 7.54]	<.0001
Vocational high school		7.00		25.05	0.23	[.15 ; .35]	<.0001
Professional diploma (CAP/BEP)		1.24		7.88	0.15	[.06 ; .37]	<.0001
Retention in grade 9		1.56		3.05	0.50	[.21 ; 1.22]	0.0165
Work		0.26		0.63	0.41	[.05 ; 3.52]	0.2472
<i>Guidance decision at the end of grade 10</i>	708		15174				
General track		82.91		63.38	2.80	[2.06 ; 3.81]	<.0001
Technological track		11.72		22.24	0.46	[.32 ; .66]	<.0001
Vocational high school		1.55		3.71	0.41	[.16 ; 1.04]	0.0027
Professional diploma (CAP/BEP)		0.14		0.7	0.03	[.01 ; 4.20]	0.0764
Retention in grade 10		3.53		9.87	0.33	[.18 ; .62]	<.0001
Work		0.14		0.09	1.65	[.07 ; 38.10]	0.472

Note : Bonferroni corrected threshold for significance is 0.0025.

Figure 2

Distribution of DNB Grades for High-IQ and Non-High-IQ Students



Note: The odd distribution with a peak at 10 is a well-known consequence of grade adjustments whose purpose is to make more students pass the exam (grade of 10).

Figure 3

Distribution of Self-Efficacy Scores in Grade 9 for High-IQ and Non-High-IQ Students

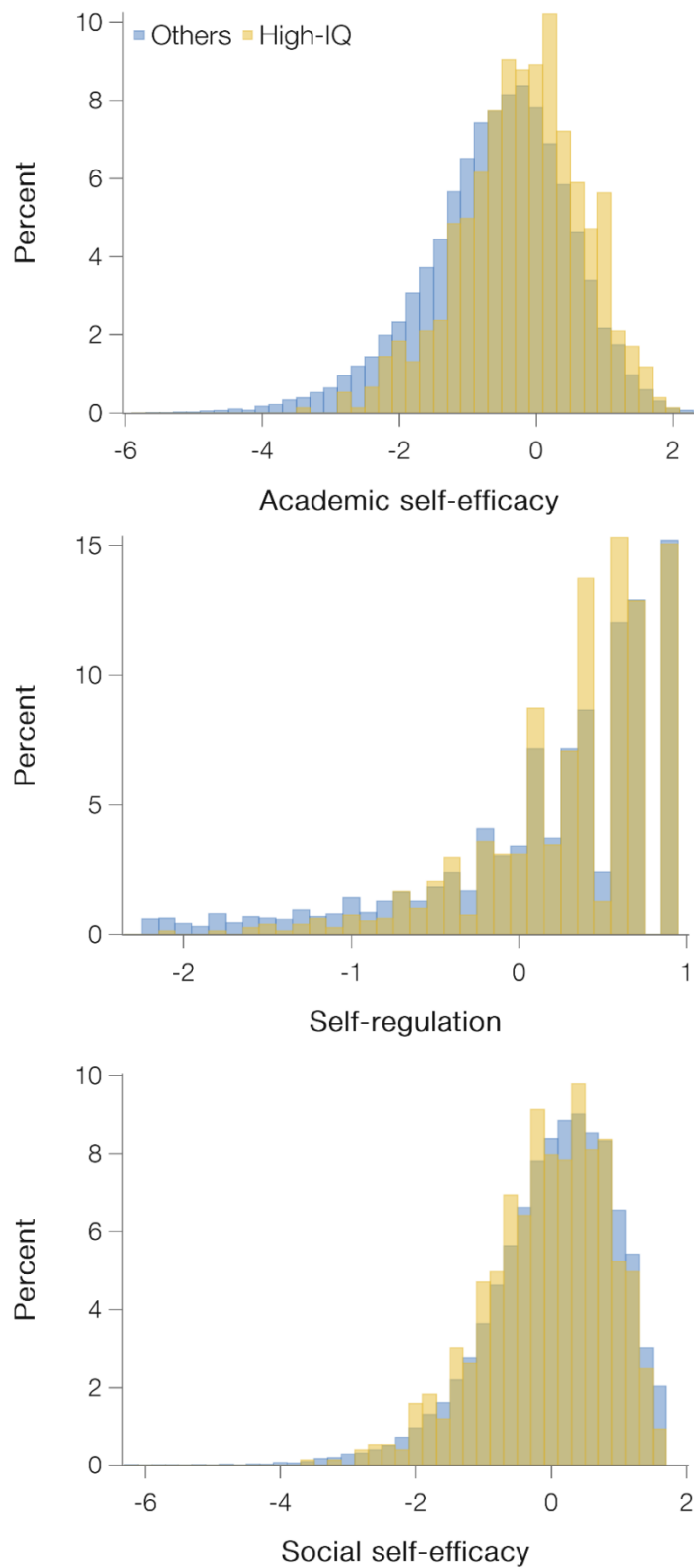
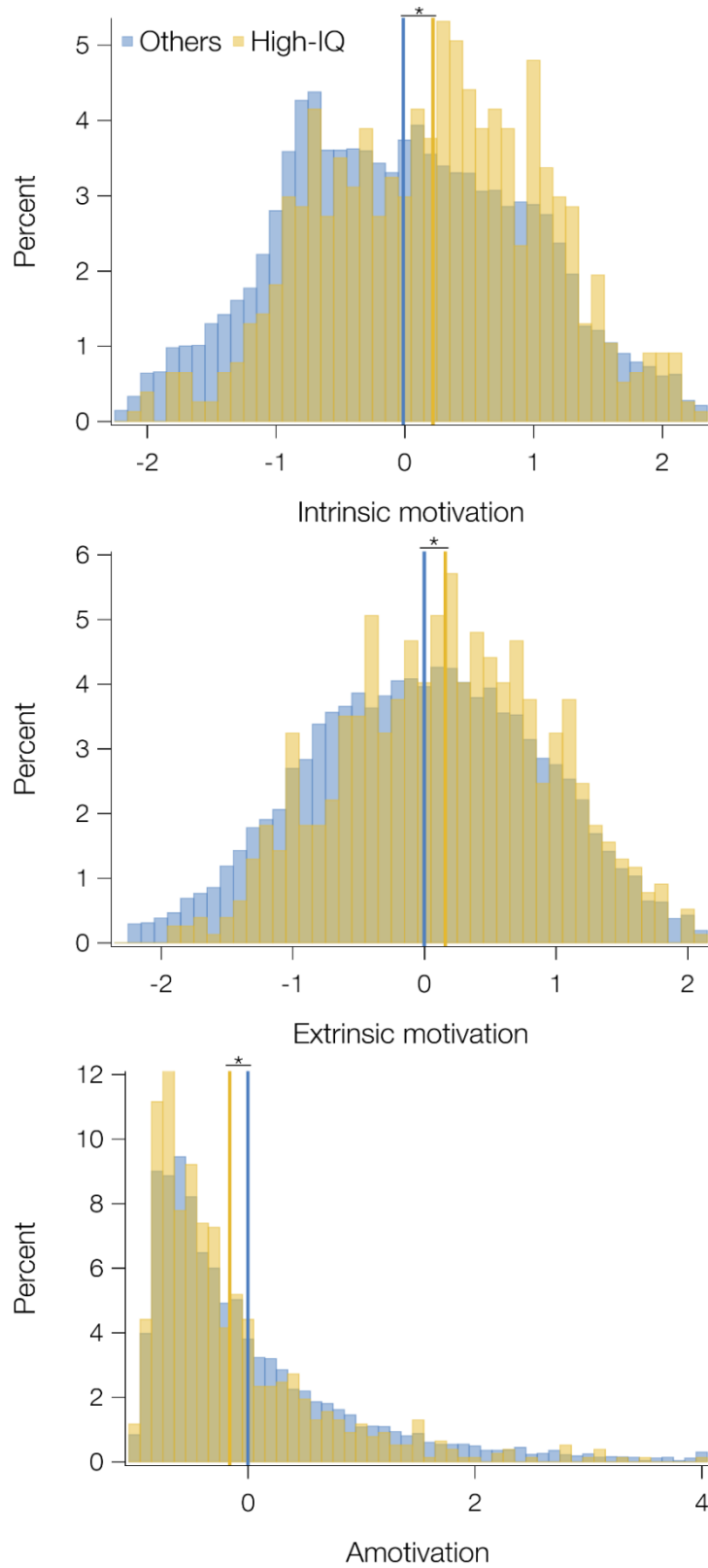


Figure 4

Distribution of Motivation Scores in Grade 9 for High-IQ and Non-High-IQ Students



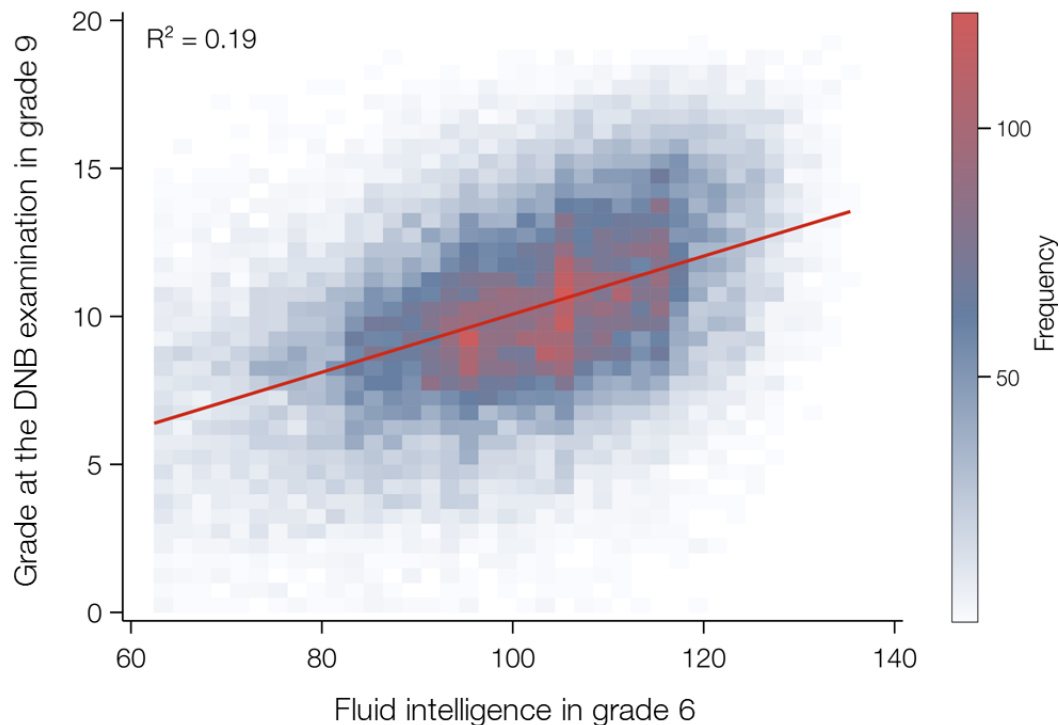
Relationship between IQ and school performance

In a second step, we tested the hypotheses that there is a positive relationship between IQ in grade 6 and achievement in grade 9 (Hypothesis d), and that this relationship holds equally in high-IQ students and in the general population (Hypothesis e). Fluid intelligence measured in grade 6 positively predicted results at the DNB in grade 9 among high-IQ students ($\beta = 0.11$, $p = 0.0058$) as well as in the whole population ($\beta = 0.10$, $p < .0001$, illustrated in **Figure 5**). This result was unchanged when we included as high-IQ all students with non-verbal IQ scores higher than the 95th percentile ($N=1602$): the regression coefficient was still positive and statistically significant ($\beta=0.15$ with $p < .0001$). However, when we increased the threshold of high-IQ to the top 1% of the non-verbal IQ distribution ($N=435$), the model yielded a still positive but not statistically significant regression coefficient for non-verbal IQ ($\beta=0.09$ with $p=0.1887$), which may result from restriction of range and the smaller number of high-IQ students in the top 1%¹⁰.

¹⁰ Our aim in this section was not particularly to examine whether the relationship between IQ and achievement was non-linear, but rather to examine whether the well-established linear relationship (also observed here), remains the same in the high-IQ range. However, for interested readers, the fit of the linear regression model with DNB as dependent variable and IQ (continuous) as independent variable ($R^2=0.194$) was very similar to the fit of a model with an additional quadratic term ($R^2=0.196$), and with both quadratic and cubic terms ($R^2=0.196$). The nonlinearity of the relationship is thus very weak.

Figure 5

Relationship Between Fluid Intelligence and Results at the DNB Examination in the Whole Population



Interactions between high-IQ, parental education and sex

Finally, we explored potential interactions between high-IQ, sex and parental education. **Table 4** presents the results of the regression of DNB Examination Grade (range: 0-20) on high-IQ, sex and parental education (henceforth, SES). Model 1 only includes main effects of these three variables, whereas Model 2 includes interactions between high-IQ and sex, and high-IQ and SES. Model 1 shows that DNB grade is positively associated with high-IQ, but also with being a girl and with parental education. In Model 2, the coefficient of the interaction parameter *High-IQ*SES* indicates the difference in DNB results associated with one additional year of parental education between high-IQ students and others. Results show that this difference is negative ($\beta = -0.11$, $p = 0.0002$), thus suggesting that high-IQ students' performance is less related to social background than their peers' (see **Figure 6**). Hence it does not seem to be the case that a low-educated family environment is more detrimental to high-IQ's performance than others'. There was no statistically significant interaction between high-IQ and sex ($p =$

0.5742), suggesting that the relationship between high-IQ and achievement is the same in both sexes: girls do not seem to outperform boys more among high-IQ students than among non-high-IQ students. However, it is worth noting that including the interactions in the regression only increased the R^2 by .04%: there was no practical gain in explained variance.

Table 4

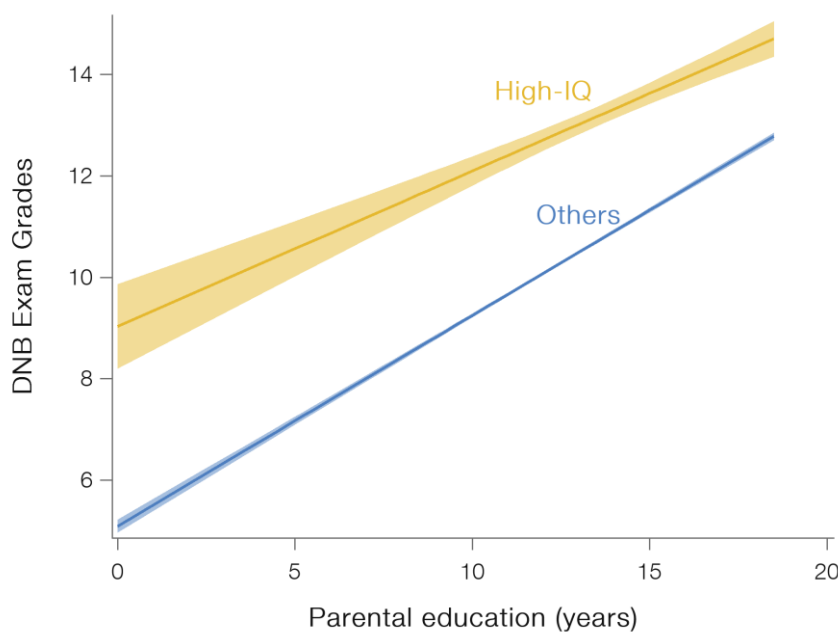
Regression of DNB Examination Grade (range : 0-20) on high-IQ, sex and parental education (SES).

Predictors	Model 1		Model 2	
	β (SD)	<i>p</i>	β (SD)	<i>p</i>
High-IQ	2.45 (0.09)	<.0001	3.86 (0.43)	<.0001
Girl	0.46 (0.04)	<.0001	0.46 (0.03)	<.0001
SES	0.41 (0.01)	<.0001	0.42 (0.01)	<.0001
High-IQ*Girl			0.10 (0.18)	0.5742
High-IQ*SES			-0.11 (0.03)	0.0002

Note : Bonferroni corrected threshold for significance is 0.01.

Figure 6

Interaction effect between High-IQ and Parental Education on DNB Exam Grades



Note: Linear fit and 95% confidence bands from the regression of DNB examination grades on parental education, high-IQ, and their interaction term.

2.2.5 Discussion

This paper aimed at assessing the differences in various school success measures between high-IQ (top 2% of non-verbal IQ distribution) and non-high-IQ students in a large representative sample of French middle school students. Our results supported the hypotheses that gifted students achieve higher academic results than other students (Hypothesis a) and drop out less (Hypothesis b), thus replicating findings from the previous literature on gifted students (Lubinski & Benbow, 2006; Roznowski et al., 2000; Terman, 1926c), and extending them to a large sample of French students from the general population. We also hypothesized that high-IQ students show higher scores in measures of self-efficacy and motivation (Hypothesis c), which was also validated in our data: high-IQ students had higher levels of academic self-efficacy and self-regulatory self-efficacy – however, their social self-efficacy was not statistically significantly different from the non-high-IQ sample. They also showed higher levels of intrinsic motivation, extrinsic motivation, and lower levels of amotivation. These results are consistent with Roznowski et al. (2000) and with Calero, García-Martín, Jiménez, Kazén, & Araque (2007), who had found that gifted children have better self-regulatory abilities than a comparable group of non-gifted children. McCoach & Siegle (2003) also reported higher academic self-perceptions in a gifted student sample compared to a general school sample.

Lastly, we hypothesized that there is a positive relationship between IQ in grade 6 and achievement in grade 9 (Hypothesis d), and that this relationship holds equally in high-IQ students and in the general population (Hypothesis e). In agreement with the past literature (Deary et al., 2007; Roth et al., 2015), we found a statistically significant positive relationship between non-verbal IQ in grade 6 and academic performance in grade 9 in the whole sample. This well-established link remained when the analysis was restricted to high-IQ students (thresholds of 2% and 5%). Furthermore, not only did the positive relationship remain, but the regression coefficient barely changed when applying different high-IQ thresholds.

As an exploratory analysis, we finally explored possible interactions between high-IQ, social background and sex. Our results suggest that academic performance is less related to social background (as measured by parental education) for high-IQ students than for their peers. Hence it does not seem to be the case that a low-educated family environment is more detrimental to high-IQ's performance than others'. On the contrary, our results support the hypothesis that high IQ buffers against the effect of low SES on

achievement. The positive association between high-IQ and achievement was similar for boys and girls. However, adding these interactions did not increase the explained variance in school performance, suggesting that their effects are without practical significance.

Overall, these results argue against the beliefs that high-IQ students are particularly at risk of school failure. This popular opinion originates from personal accounts (of failing gifted students) or from clinicians, who commonly see gifted children referred to child psychiatry services for various problems including school underachievement (Grobman, 2006; Guénoilé et al., 2013). Obvious sampling biases are inherent to clinical practice and may have contributed to spreading stereotypes regarding school failure but also psychological difficulties (Peyre et al., 2016) in gifted children. It is possible that such beliefs remain prevalent due to a lack of research based on representative populations – especially in France. The present paper intended to fill this gap, and showed that these beliefs are not supported by evidence in the French middle school context.

Thus, high-IQ children seem to be very successful in the French school environment: they earn much better grades (by about one standard deviation) at the final grade 9 national exam, and are at much lower risk of grade retention, orientation in a professional track and school drop-out. This does not imply that all of them are successful: some are not, and those ones deserve close attention, as do all pupils who do not succeed well in school. Our results do not imply either that the French education system is optimally suited to the needs of high-IQ children: they might succeed even better, or more happily, if their educational needs were better met. However, the data collected in the present study do not address the question of how to further improve the education of high-IQ children. It simply suggests that there is no major school failure problem to tackle concerning this group, and provides no basis for the necessity of a systematic screening of intellectually gifted children, as sometimes urged by associations.

Limitations

A potential limitation of our study comes from the fact that general intelligence was assessed through only one test measuring fluid intelligence (Chartier's Reasoning Test on Playing Cards). This led us to use one particular definition of giftedness, which relied exclusively on non-verbal intelligence. No measure of verbal intelligence was available in the data from the DEPP Panel 2007, therefore we could not compute a composite score more comparable to full scale IQ. However, we see no reason to suspect that the results

might have differed if a measure of verbal intelligence had been included. Indeed, including verbal intelligence should only affect our results if it affects school performance differently among non-high-IQ and high-IQ students. We do not see any reason why this should be the case.

Another limitation linked to our measure of IQ is its low stakes: students were not rewarded for performing well on it, and it was the last test in the battery. Previous research has suggested that motivation in low-stakes IQ test could be an important issue for the interpretation of the results. Compared with a rewarded condition, low-stakes would decrease IQ scores, and would do so to a larger extent for individuals with the lowest intelligence (Duckworth et al., 2011). However, the evidence is mixed, as other papers have pointed towards a positive effect of rewards on the level of effort, but that is not reflected in the IQ scores obtained (Borghans et al., 2016; Gignac, 2018). In our study, Duckworth et al. (2011)'s results could imply that we may be missing some gifted but unmotivated students in the high IQ group. Such unmotivated gifted students would be more at risk of school failure, so we might be underestimating the phenomenon. However, since the effect of motivation affects more students with low IQ, this would imply that many more students in the Others group have their IQ underestimated. Such students should nevertheless have obtained higher DNB results than expected from their underestimated IQ. Therefore, the potential confounding effect of motivation could have been to inflate DNB results of our non-gifted students group, thus reducing the difference between gifted and non-gifted students. Hence, if anything, the effect sizes would be underestimated. Furthermore, scores at our non-verbal IQ test were very weakly correlated with scores of intrinsic motivation, extrinsic motivation, and amotivation¹¹. It is likely that these motivation scores are associated to some degree with motivation to take the RCC. If this is indeed the case, then the low correlations between RCC scores and motivation scores suggest that it is implausible that our non-verbal IQ scores are confounded with motivation.

A third limitation may be that the DNB final grade, which was the main variable with which we assessed school success, followed an odd distribution: there was a peak at 10 due to grades harmonization carried out by a jury, and this peak was much more marked for non-high-IQ students than for high-IQ students. Thus, the final results of the non-high-IQ were artificially inflated compared to those of the high-IQ, which means that the

¹¹ The correlation between fluid IQ and intrinsic motivation was equal to 0.05, that between IQ and extrinsic motivation was equal to 0.03, and that between IQ and amotivation was equal to -0.07.

Chapter 2 - Intelligence and academic skills

difference between the two groups was under-estimated. However, this bias does not affect our conclusion that high-IQ students perform better than others and are less at risk of school failure (i.e. obtaining less than 10).

Lastly, our sample slightly over-represented students from schools in disadvantaged areas (*Réseau Ambition Réussite*), in which non-high-IQ students were schooled more than high-IQ students. However, correcting for this slight over-representation by using survey weights did not change our results.

Conclusion

Data from the French Depp Panel 2007 do not support the widespread belief that students with high IQ are more at risk of school failure than their peers. On the contrary, our study replicated previous findings in US samples showing that they perform better than their classmates. This result was corroborated by higher levels of motivation and self-efficacy among high-IQ students.

Chapter 3 - Early predictors of arithmetic skills

In this chapter, we investigate the early predictors of numeracy, answering the following questions: what are the early cognitive, socio-emotional and environmental predictors of academic skills; and to what extent are the effects of early environmental factors mediated through early cognitive and socio-emotional abilities (Part 3.1)? Do multiplication, addition and subtraction abilities share the same early cognitive predictors (Part 3.2)?

3.1 Cognitive and environmental predictors of problem solving skills

This section is based on the following article:

Guez, A., Peyre, H., Pinheiro-Chagas, P., Heude, B., Piazza, M., & Ramus, F. (*under review*). Early cognitive, emotional and environmental predictors of arithmetic abilities in 11.5-year-old children.

3.1.1 Abstract

Arithmetic abilities are a crucial set of skills for functioning in daily life, that are learnt and developed over many years of schooling. Understanding the set of factors that contribute to individual differences in the acquisition of these abilities is essential for designing effective teaching programs and pedagogical interventions. In the present, preregistered study, we analyzed longitudinal data from a French birth cohort (N = 343) and used structural equation modeling to assess the predictive role of a rich set of pre-school environmental, individual, cognitive and socio-emotional skills on arithmetic word problem solving at 11.5 years old. In addition, we estimated the extent to which the effects of early environmental and individual factors were mediated through pre-school cognitive and socio-emotional skills. We found that pre-school language and visuospatial abilities, but not fine-motor abilities, predicted arithmetic learning at 11.5. Beyond these early cognitive abilities, emotional symptoms (e.g. whether the child is often worried or scared, down-hearted, or nervous) negatively predicted later arithmetic performance. Parental education and sex were the most important environmental and individual predictors of arithmetic abilities, with different mechanisms at play. While the effect of parental education was partly mediated through early cognitive abilities, the effect of sex was largely direct. Our results provide insights into the cognitive, socio-emotional, and environmental factors driving the development of arithmetic abilities in children. Furthermore, they provide potential avenues of interventions to foster arithmetic abilities in children, reduce socio-economic disparities and sex differences in mathematics.

Keywords: arithmetic abilities, language abilities, visuospatial abilities, sex differences, parental education

3.1.2 Introduction

Arithmetic abilities are a set of skills key for functioning in daily and work life. As such, they are an important foundation of school programs, they influence educational outcomes and have long-term repercussions notably on the socio-economic status (SES) attained in adult life (Ritchie & Bates, 2013). Contrarily to primary quantitative abilities (the intuitive sense of numerosity and of its transformations) which are present early in life and mature with experience during infancy and early childhood, symbolic arithmetical competencies are learnt and developed with formal education, especially during primary school, where they grow considerably (Geary, 2000). Understanding the set of factors that contribute to individual differences in the acquisition of these abilities is essential for designing effective teaching programs and pedagogical interventions. Previous research has highlighted a range of early markers of arithmetic abilities, which can be divided in two categories: 1) early cognitive and socio-emotional factors, and 2) individual and environmental factors.

Early cognitive and socio-emotional predictors of arithmetic abilities

Apart from early numerical abilities (e.g., counting, number comparisons, nonverbal calculation, simple story problems, etc...), which are an obvious early marker of later arithmetic abilities (Jordan et al., 2009), several cognitive and socio-emotional traits have been found to be linked with arithmetic achievement.

First, children with higher visuospatial abilities obtain higher results in arithmetic concurrently (Hawes et al., 2019; Reuhkala, 2001; Träff et al., 2018) and are more likely to obtain higher results in the future (Yang et al., 2019; Zhang et al., 2014, 2017). Several processes may be underlying this relationship. In the first place, stronger visuospatial abilities may foster the development of an accurate linear mental representation of numbers, which in turns contributes to the variance in calculation abilities (Siegler & Booth, 2004). Moreover, complex arithmetic problems that require sequential planning may involve visuospatial abilities due to the spatial organization of calculation algorithms (Dehaene &

Cohen, 1995). Lastly, in arithmetic word problems, visuospatial abilities may support the construction of a visual schematic representation of the problem, which in turn may improve performance (Boonen et al., 2013). Clearly, there are many reasons for which those who possess strong visuospatial abilities might be well predisposed to learning math.

Second, language abilities have been found to be specific predictors of arithmetic abilities (Durand et al., 2005; Fuchs et al., 2010; Träff et al., 2018; Zhang et al., 2017). Language abilities also play multiple roles in the development of arithmetic abilities: first, when children learn to associate the rote-learned number words with the quantities they represent (Geary, 2013). In addition, simple arithmetic facts such as multiplications seem to be stored and retrieved from long-term verbal memory (Dehaene & Cohen, 1995). In the case of arithmetic word problem solving, vocabulary and language comprehension abilities are also crucial to understand the problem and translate it into an equation (Fuchs et al., 2010).

Third, higher working memory capacities are also related to better performance in arithmetic. (Alloway & Passolunghi, 2011; Andersson, 2008; Bull et al., 2008; Holmes & Adams, 2006; Passolunghi et al., 2008; Peng et al., 2016; Yang et al., 2019; Zhang et al., 2017). Working memory, including both visuospatial memory and verbal short term memory, is indeed important to hold and manipulate information while solving arithmetic problems, for example in problems which require several operations to be performed, or in multidigit operations which require carrying digits from one column to the other.

Fourth, fine motor abilities appear to be correlated with basic numerical abilities (Gashaj et al., 2019; Hawes et al., 2019) and arithmetic performance (Carlson et al., 2013; Gomez et al., 2015; Grissmer et al., 2010b; Luo et al., 2007b). This relationship may stem from the fact that children with stronger fine motor abilities have better finger-based representations of magnitudes, which may contribute to developing number sense: thus, finger gnosis has been found to predict later numerical abilities (Costa et al., 2011; Noël, 2005; Penner-Wilger et al., 2007). However, it is possible that this relationship between fine motor abilities and numerical abilities may simply be accounted for by the visuospatial components of motor tasks. Thus, fine motor tests targeting visuospatial integration (copying geometric figures) have been found to predict mathematics achievement, but not those targeting simple visual motor coordination (tracing geometric figures) (Carlson et al., 2013).

Lastly, socio-emotional skills have long been associated with general academic achievement. Indeed, children with externalizing problems (e.g. inattention-hyperactivity,

conduct disorder) may learn less in the classroom due to inattention and misbehavior, while internalizing problems (e.g. anxiety, depression) may reduce the motivation, participation and concentration of pupils in class; and both are detrimental to peer interactions (Bub et al., 2007). Mathematical achievement, in particular, has been found to be associated with both internalizing and externalizing problems (Claessens et al., 2009; Dobbs et al., 2006).

Individual and environmental predictors of arithmetic abilities, and the need to consider mediation effects

Besides early cognitive and socio-emotional factors, individual and environmental characteristics also seem to play an important part in determining academic achievement, notably in arithmetic. Male sex (Else-Quest et al., 2010; Hyde et al., 1990; Reilly et al., 2015, but see Lindberg, Hyde, Petersen, & Linn, 2010), higher socio-economic status (SES) (meta-analysis by Sirin, 2005), lower number of siblings (Siegler et al., 2012), or higher birthweight (Taylor et al., 2009) have all been positively associated with performance in mathematics.

Beyond their association with arithmetic outcomes, most of these characteristics are more broadly associated with the child's cognitive development as a whole, contributing to individual differences in cognitive and socio-emotional skills extremely early in life. In this sense, the association between individual and environmental characteristics and arithmetic abilities is likely to be at least partly mediated via early cognitive and socio-emotional factors. Hence, individual and environmental factors may be thought of as distal factors (i.e. located near the beginning of the causal chain, and acting via intermediary factors), while early cognitive and socio-emotional factors may be thought of as more proximal (intermediary) factors. Understanding these mediation relationships is essential to better disentangle and comprehend the roles of these different factors on the development of arithmetic abilities and guide future interventional research.

However, most studies on the cognitive predictors of arithmetic abilities did not explicitly model and assess these potential mediation links: individual and environmental characteristics are typically solely included as control variables. A recent notable exception is the study by Slusser, Ribner, & Shusterman (2019), which investigated in a sample of 140 American children whether the effect of parental education on early arithmetic performance was mediated through early language abilities, and found a significant mediation effect.

Chapter 3 - Early predictors of arithmetic skills

The present study uses a similar approach, but extends it by integrating a broad range of potential individual, environmental, cognitive and socio-emotional predictors in a coherent mediation model. We aim to answer the following questions: which factors measured at the end of kindergarten predict arithmetic abilities acquired by the end of primary school? More precisely, i) what are the early cognitive and socio-emotional precursors of arithmetic abilities at the end of primary school? ii) which individual and environmental factors are most crucial in influencing the development of arithmetic abilities, and iii) to what extent are their effects mediated by early cognitive and socio-emotional skills?

We tackled these questions by analyzing data from a French birth cohort study (the EDEN mother-child cohort, $N = 343$) where pre- and postnatal environmental and individual characteristics were reported by mothers; language, visuospatial, fine motor abilities, socio-emotional problems, as well as numeracy abilities, were assessed at 5.5 years; and outcome arithmetic abilities (word problem solving) were tested at 11.5 years. We hypothesized that: a) language, short-term memory, visuospatial, fine motor abilities at 5.5 years predict arithmetic abilities at 11.5 years; b) early socio-emotional skills predict additional variance in arithmetic abilities; c) early environmental and individual factors (parental education, household income, parental stimulation, number of siblings, mother's age, gestational age, maternal depression, maternal consumption of tobacco and alcohol during pregnancy, breastfeeding, birthweight, and sex) explain differences in arithmetic abilities; and d) the effects of these early environmental and individual factors are mediated through early cognitive and socio-emotional skills.

3.1.3 Method

Sample

We used data from the EDEN mother child cohort. Pregnant women seen during a prenatal visit at the departments of Obstetrics and Gynecology of the French Hospitals of Nancy and Poitiers before their twenty-fourth week of gestation were invited to participate. Exclusion criteria included a personal history of diabetes, twin pregnancy, intention to deliver outside the university hospital or to move out of the study region within the following 3 years, and inability to speak French. The participation rate among eligible women was 53%. Enrolment started in February 2003 in Poitiers and in September 2003 in

Nancy and lasted for 27 months in each center and resulted in the inclusion of 2,002 pregnant women (1,907 still in the cohort at delivery). The study was approved by the Ethical Research Committee of Bicêtre Hospital and by the Data Protection Authority (Commission Nationale de l'Informatique et des Libertés). Informed written consents were obtained from parents for themselves at the time of enrollment and for the newborn after delivery. Detailed data on children's environment and cognitive development were regularly collected from birth to 11.5 years old. At the age of 11.5 years old, 539 parents completed the questionnaires, 417 children completed the arithmetic test, and 343 (Age: $M = 11.58$, $SD = 0.52$; sex assigned at birth: 49% male children) had data meeting our inclusion criteria (listed in the **Analyses** section). The socio-demographic and individual characteristics of our working sample are displayed in **Table 1**.

Table 1
Socio-demographic and individual characteristics

Variable	N	Mean	SD	Median	Min	Max
Parental education (years)	343	14.25	2.18	14.00	10.00	17.00
Household income (k€)	343	2885.34	914.19	2837.50	975.00	5066.67
Number of siblings at 5.5 years old	343	1.23	0.80	1.00	0.00	6.00
Sex (1 = Male)	343	0.49	0.50	0.00	0.00	1.00
Gestational age (weeks)	343	39.36	1.58	40.00	31.00	42.00
Birthweight (kg)	343	3.35	0.48	3.34	1.28	5.26
Mother's age (years)	343	29.93	4.70	30.00	17.00	42.00
Alcohol consumption during pregnancy (nb drinks/week)	343	0.52	1.41	0.00	0.00	17.00
Tobacco consumption during pregnancy (1=Yes)	343	0.14	0.34	0.00	0.00	1.00
Breastfeeding duration (1=more than 3 days)	343	0.78	0.42	1.00	0.00	1.00
Cognitive stimulation (HOME score, out of 27)	335	17.42	2.13	18.00	11.00	21.00
Maternal depression episode until 5.5 years old (1=Yes)	343	0.33	0.47	0.00	0.00	1.00
Age at the last wave (years)	334	11.58	0.52	11.52	10.49	13.28
Schooled in grade 5 (1=Yes)	343	0.20	0.40	0.00	0.00	1.00
Schooled in grade 6 (1=Yes)	343	0.58	0.49	1.00	0.00	1.00
Schooled in grade 7 (1=Yes)	343	0.21	0.41	0.00	0.00	1.00

Measures

Arithmetic test at 11.5 years old. At the last collection wave, the Arithmetic subtest from the WISC-IV (Wechsler, 2003, 2005) was administered online and taken at home on the family computer, as part of a larger test battery. It consisted of 23 short arithmetic problems of increasing difficulty resembling elementary school ones. In this computerized adaptation, problems were presented both in spoken and written form; children had to solve them mentally and then type in their answer on the keyboard (in digits). The final score consisted in the sum of correct answers. Due to the fact that children were not all in the same school grade when tested (ranging from grade 5 to grade 7, see **Table 1**), we adjusted the final score for school grade (taking the residuals from the regression of the raw score on school grade). Internal consistency was good (Cronbach's α was 0.75). Examples of items:

- « *Chloé a 12 ballons et elle en vend 5. Combien de ballons lui reste-t-il ?* » (Chloé has 12 balloons and sells 5 of them. How many balloons does she still have?)
- « *Un demi-litre d'eau coule dans un seau en 1 minute. Combien faut-il de temps pour obtenir 10 litres d'eau dans le seau ?* » (Half a liter of water is poured in a bucket in 1 minute. How much time is needed to get 10 liters in the bucket?)

Cognitive tests and socio-emotional questionnaires at 5.5 years old. At 5.5 years old, children's cognitive abilities were assessed at home by a trained psychologist, by means of a range of psychometric tests. All psychometric test scores were adjusted for the child's age and birthplace. In addition, mothers filled in a questionnaire on children's mental health.

Numerical abilities. Children's numerical abilities were evaluated at 5.5 years old with the Number Knowledge Test (NKT) (Okamoto & Case, 1996), which assessed their knowledge of the number sequence as well as their ability to solve simple symbolically presented arithmetic problems. It is scored as the number of correct answers to 18 calculation exercises (counting, adding, and subtracting). Internal consistency was acceptable (Cronbach's $\alpha = 0.61$). Examples of items:

- « *Quel est le plus grand, 5 ou 4 ?* » (Which one is the largest, 5 or 4?)
- « *Combien font 8 moins 6 ?* » (How much is 8 minus 6?)

Specific cognitive abilities. Subtests from the WPPSI-III (Wechsler, 1967, 2004) and NEPSY (Kemp et al., 2001; Korkman et al., 2003) batteries were administered, as well as the Peg-moving task (Nunes et al., 2008).

- *Peg-moving task*: Children had to move five pegs, one by one, in a forward motion pattern, beginning with the peg at the side of each hand. The task started with the preferred hand and the subject had to perform three complete trials with each hand. This test is designed to measure visual-motor coordination. It is scored as the time taken by the child to move the pegs, in seconds. Internal consistency was good (Cronbach's $\alpha = 0.81$).

- *Non-word repetition (NEPSY)*: This test is scored as the number of syllables repeated correctly (out of 46 syllables in 13 non-words). It taps phonological processing (encoding and decoding) and verbal short-term memory. Internal consistency was good (Cronbach's $\alpha = 0.80$).

- *Word segment recognition (NEPSY)*: Children had to identify 14 pictures that represent words (named in the first place by the psychologist) formed from orally presented word segments. This test taps phonological processing (phonemic awareness). Internal consistency was poor¹² (Cronbach's $\alpha = 0.43$).

- *Sentence repetition (NEPSY)*: This test is scored as the number of sentences (out of 17) repeated correctly. This test is designed to measure syntactic abilities and verbal short-term memory. Internal consistency was good (Cronbach's $\alpha = 0.77$).

- *Design-copying task (NEPSY)*: Children had to copy 18 two-dimensional figures correctly (each item was rated from 0 to 4). This test taps visual perception and organization and visual-motor coordination. Internal consistency was good (Cronbach's $\alpha = 0.74$).

- *Information (WPPSI-III)*: Children had to correctly answer (verbally or by pointing) 34 questions that address a broad range of general knowledge topics (34 items). This test measures language comprehension, conceptual knowledge and verbal expressive ability.

- *Vocabulary (WPPSI-III)*: Children had to correctly define 25 words. This test is designed to measure receptive vocabulary, conceptual knowledge and verbal expressive ability.

- *Word reasoning (WPPSI-III)*: Children had to correctly identify a concept from a series of clues (28 items). This test taps language comprehension, conceptual knowledge and general reasoning ability.

¹² The fact that reliability for this measure of language was quite low is not a problem in our study since we do not use this measure alone, but as part of a latent construct for language abilities, see the Analyses section.

- *Block design (WPPSI-III)*: Children had to correctly recreate two-dimensional designs using blocks (20 items). This test is designed to measure nonverbal concept formation, visual perception and organization and visual-motor coordination.

- *Matrix reasoning (WPPSI-III)*: Children had to correctly complete 29 matrices correctly completed (29 items). This test taps nonverbal concept formation and visual perception and organization.

- *Picture concepts (WPPSI-III)*: Children had to correctly select 2 or 3 pictures with common characteristics (28 items). This test is designed to measure abstract categorical reasoning ability.

- *Coding (WPPSI-III)*: In limited time (120 seconds), children had to copy symbols, correctly pairing 2 symbols together (59 items). This test taps processing speed, visual-motor coordination and visual working memory.

The manual of WPPSI-III reports evidence of high subtest reliability (0.83 to 0.95), internal consistency, test-retest stability, and validity for all subtests (Wechsler, 1967)¹³.

Socio-emotional skills. Mothers answered the Strength and Difficulties Questionnaire (SDQ) (Goodman, 1997, 2001; Shojaei et al., 2009) when the child was 5.5 years old. The SDQ is a screening questionnaire for behavioral problems in children, which includes questions about 25 positive and negative attributes, equally divided into five subscales:

- *Emotional symptoms*: e.g. “Many worries, often seems worried”, “Nervous or clingy in new situations, easily loses confidence”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.61$).

- *Conduct problems*: e.g. “Often has temper tantrums or hot tempers”, “Often fights with other children or bullies them”. Internal consistency was good (Cronbach’s $\alpha = 0.74$).

- *Hyperactivity/inattention*: e.g. “Restless, overactive, cannot stay still for long”, “Easily distracted, concentration wanders”. Internal consistency was good (Cronbach’s $\alpha = 0.76$).

- *Peer relationship problems*: e.g. “Rather solitary, tends to play alone”, “Picked on or bullied by other children”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.50$).

- *Prosocial behavior*: e.g. “Considerate of other people’s feelings”, “Kind to younger children”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.69$).

¹³ We did not have access to the data per item for the WPPSI-III, therefore we could not compute reliability estimates for our sample.

Each item was rated by the mother using a 3-point Likert scale (1 = “Not true”, 2 = “Somewhat true”, 3 = “Certainly true”). The reliability estimates in our sample were similar to those found in a representative sample of 1,348 French children aged 6–11 years old (Shojaei et al., 2009).

Early individual and environmental factors.

Individual characteristics and birth factors. Child sex (coded as 0 for female sex and 1 for male sex), gestational age, birthweight, and the maternity hospital in which the mother was recruited and the child was born (Nancy or Poitiers), were reported at birth.

Prenatal factors. Mothers indicated their tobacco and alcohol consumption during pregnancy. Tobacco consumption was dummy coded (1 when the mother smoked and 0 otherwise) and alcohol consumption was continuous (number of glasses per week).

Family characteristics and post-natal factors. Mothers reported their age, their and the father’s level of education (we used a continuous variable equal to the average of both parents’ years of education), their household income, the child’s number of siblings, and breastfeeding duration (dummy variable equal to 1 if the child was breastfed at least 3 days, 0 otherwise). Parental cognitive stimulation of the child at home was assessed by a psychologist using three subscales of the Home Observation for the Measurement of the Environment inventory: language stimulation, academic stimulation, and variety of experimentations (27 items) (Bradley & Caldwell, 1984; Frankenburg & Coons, 1986). A HOME score was computed as the mean of all items. Lastly, mothers answered the Center for Epidemiologic Studies – Depression (CES-D) questionnaire (Radloff, 1977) (20 items) when the child was aged 4 months, 8 months, 1 year, 3 years and 5 years. Scores range from 0 to 60, with high scores indicating greater depressive symptoms. A cutoff score (16 or greater) is normally used in order to identify mothers at risk for clinical depression. We used a dummy-coded variable indicating whether the mother had at least one depressive episode between child age 4 months and 5 years.

Analyses

Our analysis plan was preregistered on OSF¹⁴. When we proceeded with the analyses, we slightly departed from the preregistration in a few respects. In addition, we ran additional specifications of the preregistered models (which we also report in the

¹⁴ https://osf.io/wxt7n/?view_only=3fa381dce2e14b1e97e8a07616e8f222

Results section). All such cases are indicated in footnotes or in the main text below. Data processing was performed with the software SAS 9.4, and statistical analyses were performed with the software Mplus 8¹⁵.

Exploratory and confirmatory factor analyses of specific cognitive abilities test scores at 5.5 years old. We first aimed to define a limited number of latent cognitive ability factors from our twelve specific cognitive ability test scores at 5.5 years, in order to subsequently use them as predictors of arithmetic abilities at 11.5 years. We ran exploratory factor analyses (EFAs) with oblimin rotation (we assumed that the factors would be correlated) and maximum likelihood estimation, varying the number of factors and checking the theoretical relevance of the suggested factor structures. We then ran confirmatory factor analyses (CFAs) based on the results of the EFAs, checking the fit indices (RMSEA, CFI and TLI). Factor analyses were conducted on the entire sample of children present at the 5.5 year-old wave who had available psychometric test scores (N=1,129)¹⁶.

We then turn to the models answering our research questions. For all the following models, we excluded from the analysis children who had missing data on the arithmetic test at 11.5 years old, on the NKT, on more than 6 psychometric tests, on the SDQ, or on both measures of the socio-economic environment (parental education and household income) (remaining N=343; see **Figure A1** in Appendix). The remaining missing data was treated using full information maximum likelihood (FIML) estimation.

Models A1 and A2: Cognitive and socio-emotional precursors of arithmetic abilities at 11.5 years old. Our first research question regarded the early cognitive and socio-emotional precursors of arithmetic abilities at 11.5 years old. Based on the factor structure resulting from the factor analyses, we ran a series of structural equation models with the arithmetic score at 11.5 years old as our outcome variable. We introduced our two blocks of predictors in a sequential manner (first the latent cognitive factors – Model A1, then adding socio-emotional difficulties – Model A2), with and without including the NKT

¹⁵ Both our intended SAS and Mplus scripts were posted on OSF along with the pre-registration. Some changes were made in both scripts. Readers can find our initial preregistered scripts as well as our final ones on OSF (https://osf.io/wxt7n/?view_only=3fa381dce2e14b1e97e8a07616e8f222), with differences flagged.

¹⁶ In the preregistration, we had planned to run the factor analyses on the final sample (children with data at 11.5 years old). However, we subsequently realized that running them on the largest possible sample (all children with data at 5.5 years old) would yield a more reliable factor structure.

as an additional predictor (Models A1' and A2') in order to understand to what extent early cognitive and socio-emotional skills predicted later arithmetic abilities independently from the child's baseline mathematics level at 5.5 years old. We checked modification indices to test for significant correlations between predictors, which were declared in the model when necessary. The models' goodness of fit was examined using the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean squared error of approximation (RMSEA). CFI and TLI values greater than 0.95 and values of RMSEA less than 0.06 were used as cut-offs. We only interpreted standardized coefficients in order to compare the effect sizes between different predictors. The significance of coefficients was assessed with two-sided t-tests. We used the Benjamini-Hochberg False Discovery Rate (FDR) and a false discovery threshold $q=0.05$ to take into account the number of regression coefficients estimated in each model.

Model B: Individual and environmental factors predicting arithmetic abilities at 11.5 years old. Our second research question concerned the pre- and post-natal individual and environmental predictors of arithmetic abilities. We ran a multivariate regression with all such factors as independent variables, and arithmetic abilities as our dependent variable. Similarly to the previous models, we used standardized coefficients, and assessed significance with the FDR ($q=0.05$).

Models C: Mediation models. Lastly, we aimed at exploring how the associations between early individual and environmental factors and arithmetic were mediated through early cognitive and socio-emotional skills. Running a mediation model with the SEM framework involves (i) regressing the arithmetic score on the individual and environmental (distal) predictors (*direct paths*) and on the cognitive and socio-emotional mediators, (ii) simultaneously regressing the mediators on the distal predictors, and (iii) estimating the significance of the *indirect paths* (i.e., associations between distal predictors and arithmetic which are mediated through cognitive and socio-emotional skills). In these mediation models, the estimation of the size of the standardized mediation effects was the main purpose, so we did not control for multiple testing. As preregistered, we only included in the mediation model predictors which were significantly associated with arithmetic performance in the multivariate regression (Models C and C' – without and with the NKT as an additional predictor). Again, we checked modification indices to test for significant

correlations between predictors, which were declared in the model when necessary, and checked goodness-of-fit using the CFI, TLI and RMSEA. Correlations between all mediators were included in order to account for any unmodeled sources of covariation among mediators, and bias-corrected bootstrapped confidence intervals with 10,000 resamples were derived in order to account for the likely non-normality of mediated effects (product of estimates), following recommendations by Preacher & Hayes (2008).

3.1.4 Results

We present here the main results of our analyses. Full Mplus outputs and Excel sheets are available on OSF for further information. Descriptive statistics for arithmetic, cognitive and socio-emotional abilities are presented in **Table 2**.

Table 2

Descriptive statistics for arithmetic, cognitive and socio-emotional abilities

Variable	N	Mean	SD	Median	Min	Max
Arithmetic at 11.5 years old (WISC-IV, out of 23)	343	12.45	3.46	13.00	2.00	21.00
NKT (out of 18)	343	12.45	3.42	13.00	4.00	18.00
Peg moving task (time in seconds)	341	27.66	4.70	27.00	20.00	46.00
Nonword repetition (NEPSY, out of 46)	342	29.43	7.75	30.00	5.00	45.00
Word-segment recognition (NEPSY, out of 14)	338	11.00	1.78	11.00	4.20	14.00
Sentence repetition (NEPSY, out of 34)	328	16.47	3.99	16.00	7.00	28.00
Design copying (NEPSY, out of 72)	338	52.83	7.23	52.80	31.20	69.60
Information (WPPSI-III, out of 34)	343	25.75	2.66	26.00	17.00	31.00
Vocabulary (WPPSI-III, out of 43)	342	24.60	5.51	24.00	11.00	40.00
Word reasoning (WPPSI-III, out of 28)	342	17.06	4.28	18.00	0.00	27.00
Block design (WPPSI-III, out of 40)	341	29.12	3.69	30.00	20.00	40.00
Matrix reasoning (WPPSI-III, out of 29)	343	16.28	3.97	16.00	6.00	28.00
Picture concepts (WPPSI-III, out of 28)	342	14.92	3.79	15.00	5.00	24.00
Coding (WPPSI-III, out of 65)	339	30.83	9.13	30.00	11.00	59.00
Emotional symptoms (SDQ, out of 15)	343	2.20	1.87	2.00	0.00	8.00
Conduct problems (SDQ, out of 15)	343	2.17	1.89	2.00	0.00	8.00
Hyperactivity/inattention (SDQ, out of 15)	343	2.79	2.38	2.00	0.00	10.00
Peer relationship problems (SDQ, out of 15)	343	1.08	1.22	1.00	0.00	6.00
Prosocial behavior (SDQ, out of 15)	341	8.43	1.69	9.00	3.00	10.00

Factor structure of cognitive abilities at 5.5 years old

Results from the EFA with 4 factors are displayed in **Table 3**¹⁷. The first factor loaded mostly on non-word repetition and sentence repetition, both tapping verbal short-term memory. The second factor loaded on tests assessing verbal intelligence (information, vocabulary, word reasoning – and word segment repetition and picture concepts to a lower extent). The third factor loaded on visuospatial tests (design copying, block design, matrices, and picture concepts to a lower extent). The fourth factor was less legible, loading on coding, peg-moving task, and to a lower extent word-segment recognition, design copying and picture concepts. The stronger weights on coding and the peg-moving task suggest a factor reflecting mostly processing speed and visual-motor coordination. We based our CFA on the factor structure that emerged from this EFA, setting all non-significant loadings in the EFA to 0 in the CFA, with a few changes. First, given the strong correlation ($r=0.73$) between the first and second factors from the EFA (reflecting verbal short-term memory and verbal intelligence, respectively), we decided to create global language abilities latent variable loading on nonword repetition and sentence repetition, in addition to information, vocabulary, word reasoning, word-segment recognition and picture concepts (hence choosing a 3-factors structure). Indeed, this very high correlation would have prevented us from disentangling the effect of verbal short-term memory from that of language abilities on later arithmetic abilities. Second, we removed the word-segment recognition task and the picture concept task from the last factor, keeping coding, the peg-moving task, and design copying in order to have a latent factor reflecting visual-motor coordination (hereafter named Fine-motor abilities). A correlation between nonword repetition and sentence repetition was added to the model after checking modification indices (M.I.= 177.501). Results from the CFA are illustrated in **Figure 1**.

¹⁷ In the preregistration, we had planned to run five EFAs, with 4 to 8 factors extracted. However, the rotation algorithms could not converge with more than 4 factors.

Table 3

Results from an Exploratory Factor Analysis of specific cognitive abilities test scores at 5.5 years old

	EFA with 4 factors			
	F1	F2	F3	F4
Test score				
Peg-moving task	0.05	0.05	0.14	0.29*
Nonword repetition	0.82*	-0.06*	-0.03	0.05
Word-segment recognition	0.13*	0.17*	0.01	0.18*
Sentence repetition	0.72*	0.15	0.06	-0.06*
Design copying	0.10*	-0.10*	0.54*	0.13*
Information	0.11*	0.70*	0.05	0.00
Vocabulary	0.02	0.74*	-0.06	0.02
Word reasoning	-0.02	0.87*	0.02	0.01
Block design	0.00	0.01	0.73*	-0.01
Matrices	-0.04	0.07	0.59*	-0.04
Picture concepts	-0.04	0.25*	0.27*	0.13*
Coding	-0.01	0.02	0.00	0.73*
Factor correlations				
	1.00			
	0.74*	1.00		
	0.42*	0.56*	1.00	
	0.22*	0.23*	0.44*	1.00
Fit indices				
RMSEA	0.013 [95% CI : 0.00, 0.029]			
CFI	0.999			
TLI	0.997			

Note: * denotes $p < 0.05$; EFAs with oblimin rotation

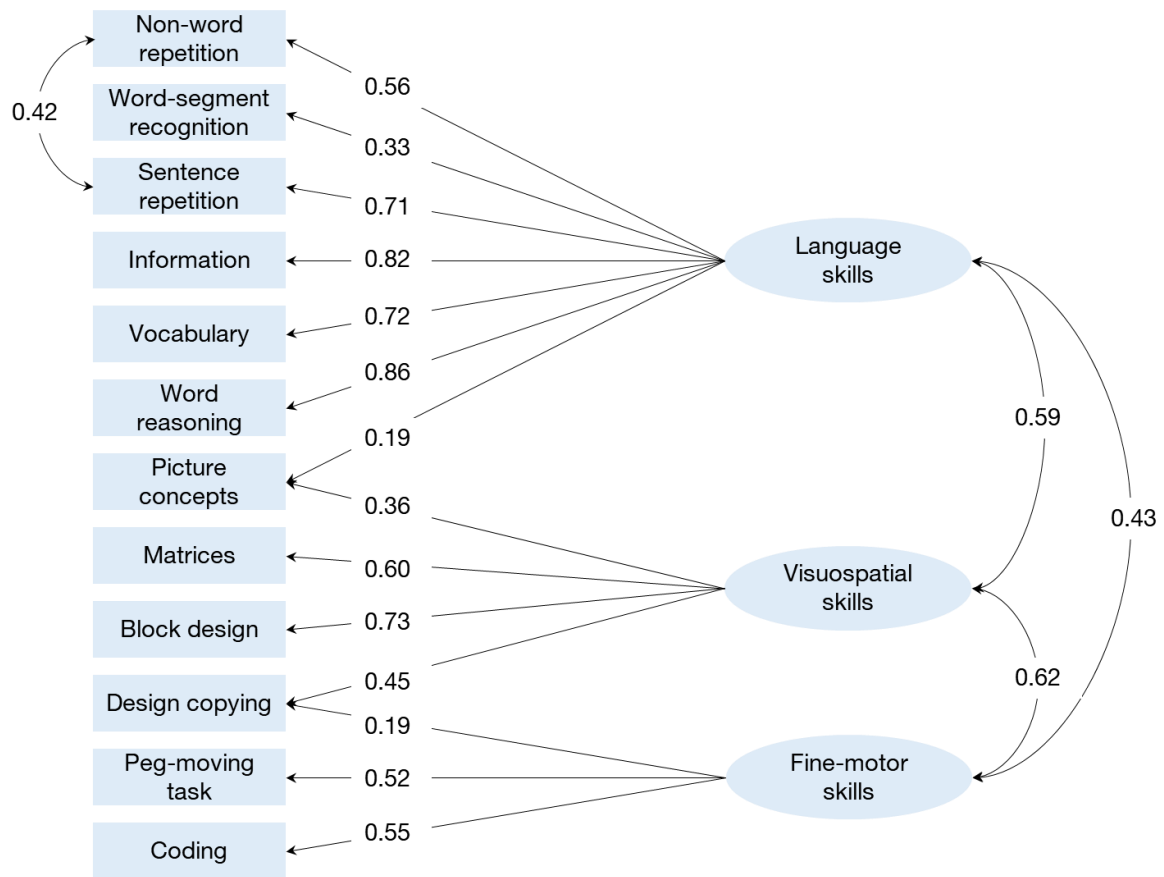


Figure 1: Confirmatory Factor Analysis of cognitive test scores at 5.5 years old. Coefficients are standardized. CFI=0.988, TLI=0.984, RMSEA [95% CI]=0.030 [0.021, 0.038]. All coefficients are significantly different from zero at the 1% level of significance.

Early cognitive and socio-emotional predictors of arithmetic abilities at 11.5 years old

Results from the longitudinal SEMs with cognitive and socio-emotional skills at 5.5 years old as predictors of arithmetic abilities at 11.5 years old are presented in **Table 4**. Language and visuospatial abilities significantly predicted future arithmetic abilities (standardized β between 0.19 and 0.22 for language abilities across all models, and between 0.29 and 0.35 for visuospatial abilities). Thus, an increase of one standard deviation in language abilities and visuospatial abilities predicted an increase in the arithmetic score of about 0.2 standard deviation and 0.3 standard deviation, respectively. On the contrary, the standardized regression coefficient for fine motor abilities was negative, rather small (less than -0.1), and non-significant. Adding in the socio-emotional difficulties scores (models

Chapter 3 - Early predictors of arithmetic skills

A2 and A2') degraded the fit of the model, which was nevertheless still good. These variables did not explain additional variance in later arithmetic abilities (22% in both cases). In fact, the emotional symptoms subscore was the only subscore which significantly (and negatively) predicted later arithmetic performance (standardized β between -0.11 and -0.12, $p < 0.05$, however it did not reach significance after FDR correction). Thus, children with higher levels of emotional symptoms tended to have lower scores in arithmetic. Lastly, and most surprisingly, the Number Knowledge Test did not significantly predict future arithmetic abilities and had a very small coefficient (its standardized β lying between 0.05 and 0.06). Since the NKT contains some items which reflect visual and verbal counting and not only numerical capacities (7 items), we checked whether removing these items from the NKT total score would change the results (details of the items removed and the Mplus outputs are available on OSF). However, the results remained unaltered after doing so. A follow-up univariate regression indicated that an increase in one standard deviation in the NKT score was associated with an increase of 0.29 standard deviation in the arithmetic score ($p < 0.001$, $R^2 = 0.085$), an effect size comparable to those of language and visuospatial abilities in the multivariate models.

Early individual and environmental predictors of arithmetic abilities at 11.5 years old

Results from the multivariate regression of the arithmetic score at 11.5 years old on individual and environmental predictors are displayed in **Table 5**. The child's sex and the level of parental education were the most important predictors of arithmetic abilities; they were also the only predictors with a significant effect. Thus, independently from other included environmental characteristics, boys had arithmetic scores on average 0.48 standard deviations higher than girls; and an increase of one standard deviation in parental level of education increased arithmetic performance by 0.22 standard deviation. Individual and environmental predictors explained about 13% of the variance in the arithmetic score ($R\text{-square} = 0.129$).

Table 4
Cognitive predictors of arithmetic abilities at 11.5 years old

	Model	A1	Model	A1'	Model	A2	Model	A2'
	β (S.D.)		β (S.D.)		β (S.D.)		β (S.D.)	
Predictors at 5.5 years old								
Number Knowledge Test			0.05 (0.06)				0.06 (0.06)	
<i>Latent cognitive skills</i>								
Language skills	0.21* (0.07)		0.19* (0.07)		0.22* (0.07)		0.19* (0.07)	
Visual-spatial skills	0.36* (0.10)		0.35* (0.10)		0.32* (0.10)		0.30* (0.10)	
Fine-motor skills	-0.12 (0.11)		-0.12 (0.11)		-0.10 (0.11)		-0.09 (0.11)	
<i>Behavioral/emotional skills</i>								
Emotional symptoms					-0.11† (0.05)		-0.12† (0.05)	
Conduct problems					0.03 (0.06)		0.03 (0.05)	
Hyperactivity/inattention					-0.05 (0.06)		-0.05 (0.05)	
Peer relationship problems					0.07 (0.05)		0.08 (0.05)	
Prosocial behavior					0.05 (0.06)		0.05 (0.06)	
Fit indices								
CFI	0.991		0.991		0.961		0.964	
TLI	0.988		0.988		0.954		0.958	
	0.020		0.020		0.031		0.029	
RMSEA	[0.000;0.039]		[0.000;0.037]		[0.016;0.042]		[0.015;0.041]	
R-square	0.223		0.224		0.218		0.221	

Note: Adjusted on recruitment centre. Standardized coefficients are reported. * denotes significance after FDR correction with $q=0.05$ and the number of tests for each model equal to the number of predictors of arithmetic in the model. † denotes coefficients with $p<0.05$ but which did not reach significance after FDR correction. In model A1' and A2', the NKT score and the latent variables Language and Visual-spatial skills were allowed to covary after checking modification indices.

Table 5*Environmental predictors of arithmetic skills at 11.5 years old*

	Standardized β	S.E.	p-value
Predictors			
Number of siblings	-0.07	0.05	0.197
Sex (Male)	0.48*	0.10	0.000
Gestational age	0.04	0.06	0.555
Birthweight	-0.02	0.06	0.786
Age of mother	0.08	0.06	0.162
Parental education	0.22*	0.07	0.001
Alcohol intake	-0.09	0.05	0.075
Tobacco intake	0.10	0.15	0.517
Household income	-0.04	0.07	0.553
Breastfeeding	-0.08	0.13	0.547
Cognitive stimulation	0.02	0.06	0.699
Maternal depression	-0.17	0.11	0.134
Center	0.14	0.12	0.228

Note: * denotes significance after FDR correction with $q=0.05$ and the number of tests equal to the number of predictors in the model. R-square=0.129

Mediation of parental education and sex effects through early cognitive and socio-emotional skills

Results from the models including sex and parental education (the only significant predictors in the multivariate regression from the last step) as predictors and early cognitive and socio-emotional skills as mediators are presented in **Table 6**, and illustrated in **Figure 2**. A direct effect refers to an association between the predictor and the outcome which is not explained by the association between that predictor and the mediators; while an indirect effect designates an association which is mediated through a given mediator. The total indirect effect is the sum of all mediated effects; and the total effect is the sum of indirect and direct effects. Sex had a large direct effect on arithmetic abilities at 11.5 years old (standardized direct effect = 0.45, 95% bias-corrected bootstrapped C.I. = [0.03, 0.83]). In fact, 85% of the total effect of sex was direct, with the remaining total indirect effect being comparatively quite small and not significantly different from 0 (0.08 [-0.25, 0.50]). The

largest indirect effect of sex was that through visuospatial abilities (0.07 [-0.02, 0.33]), although this did not reach significance. In contrast, the total effect of parental education was 40% direct (0.08 [-0.05, 0.21]) and 60% mediated through early cognitive abilities (0.11 [0.03, 0.23]) – however, contrary to the total indirect effect, the direct effect did not reach significance. The largest and only significant indirect effect of parental education was through language abilities (0.06 [0.01, 0.12]). Its mediated effect through visuospatial abilities was of the same magnitude, but not significantly different from zero (0.06 [-0.03, 0.19]). Adding in the Number Knowledge Test as a mediator barely changed the results (Model C’).

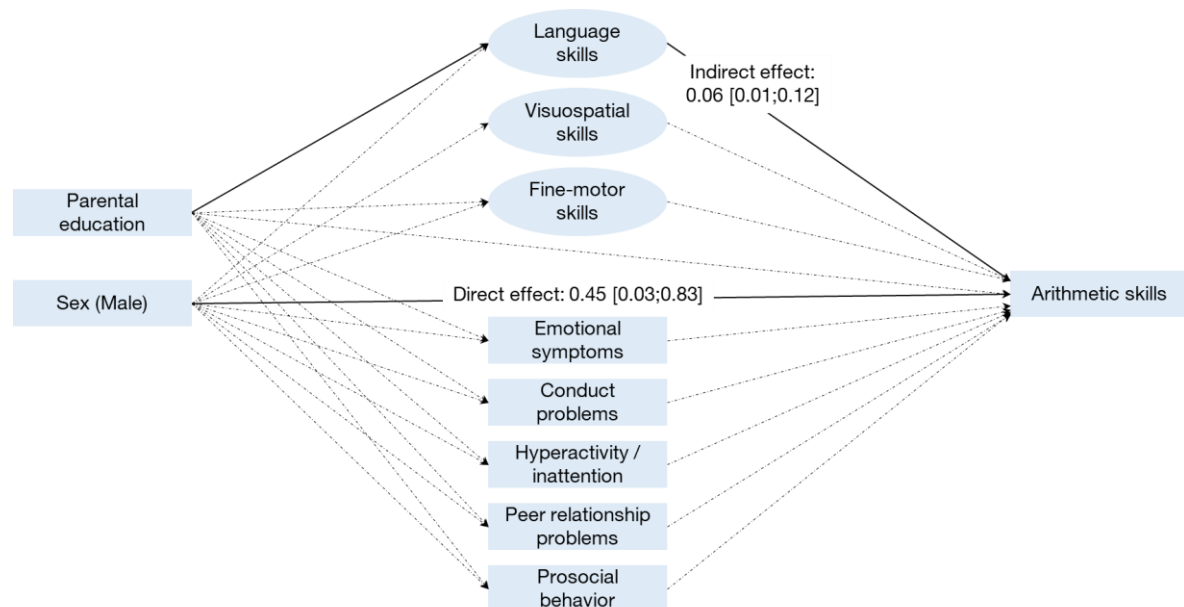


Figure 2: Mediation model (Model C). Adjusted on recruitment center. Coefficients are standardized; in brackets are the corresponding bias-corrected bootstrapped confidence intervals (BCCI) with 10,000 resamples. Only estimates of significant paths are displayed.

Table 6

Mediation effects of sex and parental education through cognitive abilities at 5.5 years on arithmetic abilities at 11.5 years old

	Model C		Model C'	
	Sex (Male) β [95% B.C.C.I.]	Parental education β [95% B.C.C.I.]	Sex (Male) β [95% B.C.C.I.]	Parental education β [95% B.C.C.I.]
Parameters				
Total	0.53 [0.33;0.71]	0.20 [0.09;0.30]	0.53 [0.33;0.71]	0.20 [0.09;0.30]
Total indirect	0.08 [-0.25;0.50]	0.11 [0.03;0.23]	0.08 [-0.12;0.37]	0.11 [0.02;0.23]
<i>Direct effect</i>	0.45 [0.03;0.83]	0.08 [-0.05;0.21]	0.45 [0.14;0.71]	0.09 [-0.05;0.21]
<i>Indirect effect through:</i>				
Number Knowledge Test			0.01 [-0.02;0.06]	0.01 [-0.02;0.04]
Language skills	0.03 [-0.01;0.11]	0.06 [0.01;0.12]	0.03 [-0.01;0.11]	0.05 [0.01;0.12]
Visual-spatial skills	0.07 [-0.02;0.33]	0.06 [-0.03;0.19]	0.07 [-0.01;0.22]	0.06 [-0.04;0.19]
Fine-motor skills	-0.00 [-0.27;0.29]	0.00 [-0.04;0.04]	-0.00 [-0.11;0.26]	0.00 [-0.04;0.05]
Emotional symptoms	0.01 [-0.01;0.06]	-0.01 [-0.03;0.00]	0.01 [-0.01;0.05]	-0.01 [-0.03;0.00]
Conduct problems	0.00 [-0.03;0.05]	0.00 [-0.01;0.02]	0.00 [-0.03;0.05]	0.00 [-0.01;0.02]
Hyperactivity/inattention	-0.01 [-0.07;0.02]	0.01 [-0.01;0.04]	-0.01 [-0.06;0.02]	0.01 [-0.01;0.02]
Peer relationship problems	0.01 [-0.00;0.05]	-0.01 [-0.03;0.00]	0.01 [-0.00;0.05]	-0.01 [-0.03;0.00]
Prosocial behavior	-0.03 [-0.10;0.01]	0.00 [-0.01;0.01]	-0.03 [-0.10;0.01]	0.00 [-0.01;0.01]
Fit indices				
RMSEA [95% CI]	0.036 [0.024;0.047]		0.035 [0.023;0.046]	
CFI	0.958		0.962	
TLI	0.933		0.936	
R-square	0.273		0.272	

Note: Standardized coefficients and 95% bias-corrected bootstrapped confidence intervals (B.C.C.I.s) with 10,000 resamples are reported. Significant coefficients are indicated in bold. In both models, all mediators were allowed to covary.

3.1.5 Discussion

This study aimed at providing a better understanding of the early cognitive, socio-emotional, individual and environmental precursors of arithmetic abilities at the end of primary school. We explored the role of a broad range of such predictors, and integrated them in a mediation framework where individual and environmental characteristics were modeled as distal factors, and cognitive and socio-emotional skills as more proximal factors.

Our results provide support for the role of language and visuospatial abilities as important building blocks of arithmetic word problem solving abilities. Indeed, and as hypothesized, better language and visuospatial abilities at 5.5 years old positively predicted arithmetic performance at 11.5 years old. In contrast, and contrary to our hypothesis, fine motor abilities had no significant predictive power above language and visuospatial abilities. This finding contradicts previous results which had found an association between fine motor abilities and arithmetic abilities after controlling for IQ (Carlson et al., 2013) or for earlier reading and mathematics scores (Grissmer et al., 2010b). It suggests that the association between fine motor abilities and arithmetic is confounded with the effect of visuospatial abilities (and perhaps language), which is in line with results showing that visuospatial integration but not visual motor coordination predicts performance in mathematics (Carlson et al., 2013). Lastly, we were not able to disentangle verbal working memory from verbal intelligence, both being reflected in our latent language factor – therefore, we could not directly test the hypothesized predictive role of working memory in this study.

Among socio-emotional skills, emotional symptoms (e.g. whether the child is often worried or scared, down-hearted, or nervous) were the only significant predictor of later arithmetic abilities, such that children suffering more from such symptoms at 5.5 years old obtained lower results in arithmetic at 11.5 years old. Scoring high on the emotional symptoms subscale may have reflected a higher propensity to math anxiety (Wu et al., 2014), which is itself associated with lower mathematics performance (Wu et al., 2012, 2014) – and especially so in online timed tests like the one we used (Ashcraft, 2002). More research is needed to fully understand the mechanisms underlying this result.

Surprisingly, early numerical abilities, as measured by the NKT test, did not predict additional variance in the arithmetic test above language and visuospatial abilities (despite a raw association of $\beta = 0.29$). This would seem to suggest that what the NKT captures of

numerical ability is confounded with language and visuospatial abilities. This may be due to the fact that the NKT does not include tests of the approximate number sense, which is the one numerical ability that does not seem to be entirely confounded with language and visuospatial skills (Libertus et al., 2013; Mazzocco et al., 2011; Starr et al., 2013; but see Chinello et al., 2013). Further research using more specific measures of early numerical abilities might be able to uncover a third, independent cognitive contributor to later arithmetic skill.

Among the explored distal environmental and individual predictors, the child's sex and the level of parental education were the only factors to have a significant effect on arithmetic abilities at 11.5 years old. Thus, boys outperformed girls, as previously found in meta-analyses (Else-Quest et al., 2010; Hyde et al., 1990; Reilly et al., 2015) and in the PISA (Programme for International Student Assessment) studies, where at age 15, boys outperformed girls in most participating countries including France (OECD, 2015). Similarly, children from higher-education families outperformed those from lower-education backgrounds, which is in line with meta-analytic results showing the strong correlation between mathematic achievement and SES (S. R. Sirin, 2005). This relationship was also previously found in France in a different cohort, in which parental education was one of the strongest environmental predictors of achievement and of its progression during middle school (Guez et al., 2018).

Integrating these distal and proximal predictors in a mediation model allowed us to understand the potential pathways through which sex and parental education ultimately influence a child's performance in arithmetic. We found a largely direct effect of sex on arithmetic performance at 11.5, suggesting that the observed difference in arithmetic between girls and boys at the end of primary school is mostly unexplained by early differences in language and visuospatial abilities before entry in primary school. This finding is consistent with previous results in large US samples, where the gender gap in mathematics was non-existent in kindergarten (Kersey et al., 2018) and appeared and developed during elementary school (Robinson & Lubienski, 2011). Two different (not mutually exclusive) causes could be underlying this fact: on the one hand, it may be that biological differences between girls and boys develop from around 6 years of age; on the other hand, it may be that unmeasured environmental or individual factors which are associated with arithmetic abilities affect girls and boys differently during primary school – for instance, one of these factors may be teachers' attitude: primary school female teacher's math anxiety has been shown to reinforce female pupils' stereotype regarding

girls' lower abilities at math, and to negatively affect their math achievement (Beilock et al., 2010). While we are unable to adjudicate between these two possible explanations in this study, international comparisons seem to favor an environmental interpretation. Indeed, results from the TIMSS 2015 (Trends in International Mathematics and Science Study) studies (Mullis et al., 2016) – which confirm the existence of a gender gap in mathematics in 4th grade in France – indicate that the international average gap was null, with 23 countries out of 49 showing no difference, 18 countries showing a gap in favor of boys, and 8 in favor of girls in 4th grade. In 8th grade the picture was similar, with 26 countries out of 39 showing no difference, 7 showing a gap favor of girls and 6 in favor of boys. In PISA, where children were assessed at 15 years old, the international average gap was in favor of boys, with 37 countries out of 64 showing a gap in favor of boys, 5 countries in favor of girls, and 22 countries showing no difference. These results thus highlight the extent to which environmental factors can alter the observed gender gap, even if they do not provide any definite answer to the present question. Hence, this calls for further research on the emergence of sex differences in mathematics during primary school in France.

Lastly, the effect of parental education was interestingly half direct, half indirect (although the direct effect did not appear to be significant), with the indirect effect being equally divided between early language abilities and early visuospatial abilities (the latter non-significant). These results thus suggest that parental education exerts effects on academic achievement at multiple stages, including early effects on cognitive skills that later affect the acquisition of arithmetic abilities, and later, enduring effects on achievement, for instance in the form of help with homework or resources that bolster academic achievement. The significant indirect effect through early language that we found corroborates findings in an American sample by Slusser et al. (2019). Such an effect might be related to the finding that, at higher levels of SES, higher math skill is associated with greater recruitment of brain areas involved in language processing (Demir et al., 2015).

Limitations and strengths

Our results should be interpreted in the light of the following limitations. Firstly, we could only include a limited range of latent cognitive predictors of arithmetic at 5.5 years old: other known predictors of arithmetic abilities could thus not be included, such as executive functions in general (Bull et al., 2008; Bull & Scerif, 2001; Espy et al., 2004;

Gashaj et al., 2019; LeFevre et al., 2013, but see Hawes et al., 2019), and particularly working memory (Alloway & Passolunghi, 2011; Andersson, 2008; Bull et al., 2008; Holmes & Adams, 2006; Passolunghi et al., 2008; Peng et al., 2016; Yang et al., 2019; Zhang et al., 2017). The absence of these factors may have biased the effects of other cognitive factors upwards. Related to this limit, we must press the fact that in spite of our efforts to include a broad range of predictors of arithmetic abilities such that the influence of each of these predictors is as little as possible confounded with that of other factors, our study remains correlational and predictive in nature. Lastly, selective attrition between inclusion and 11.5 years old limits the external validity of our results. Indeed, children included in our study are from higher socio-economic backgrounds and have higher cognitive abilities than those present at the 5.5 years wave who could not be followed-up and included in our sample (see **Table A1** in Appendix). We should also emphasize that, although we have termed some of the factors (such as parental education) as “environmental”, the lack of control for genetic factors makes it impossible to determine to what extent these effects are truly environmental, and to what extent they are genetically mediated (for example, see Krapohl & Plomin, 2016).

In spite of these limitations, these study contributes to the literature with a number of strengths, including a relatively large sample size ($N = 343$) compared to usual studies in the domain, a longitudinal design spanning the entire period from birth to the end of primary school, a rich set of predictors – individual, environmental, cognitive and socio-emotional, and a mediation design which allowed us to bring new insights into the potential mechanisms leading to the formation of individual differences in arithmetic abilities.

Conclusions and practical implications

Our results support the role of language and visuospatial abilities as important pathways to arithmetic learning and exhibit the detrimental influence of negative emotional symptoms on later arithmetic performance. This suggests the importance of insisting on language and visuospatial abilities learning, and of better considering and integrating emotional regulation abilities as well as mental health issues in the classroom. Besides, our findings confirm the well-known role of parental education as a crucial determinant of arithmetic learning and highlight its potential mechanisms, in particular through early language abilities. Thus, tackling early difficulties in language may also help dampening SES gaps in mathematics. Lastly, our results display important sex differences in outcome

arithmetic results, which were not explained by early cognitive abilities. This suggests that interventions aiming at reducing the gender gap in mathematics should focus on primary school.

3.1.6 Appendix

Figure A1

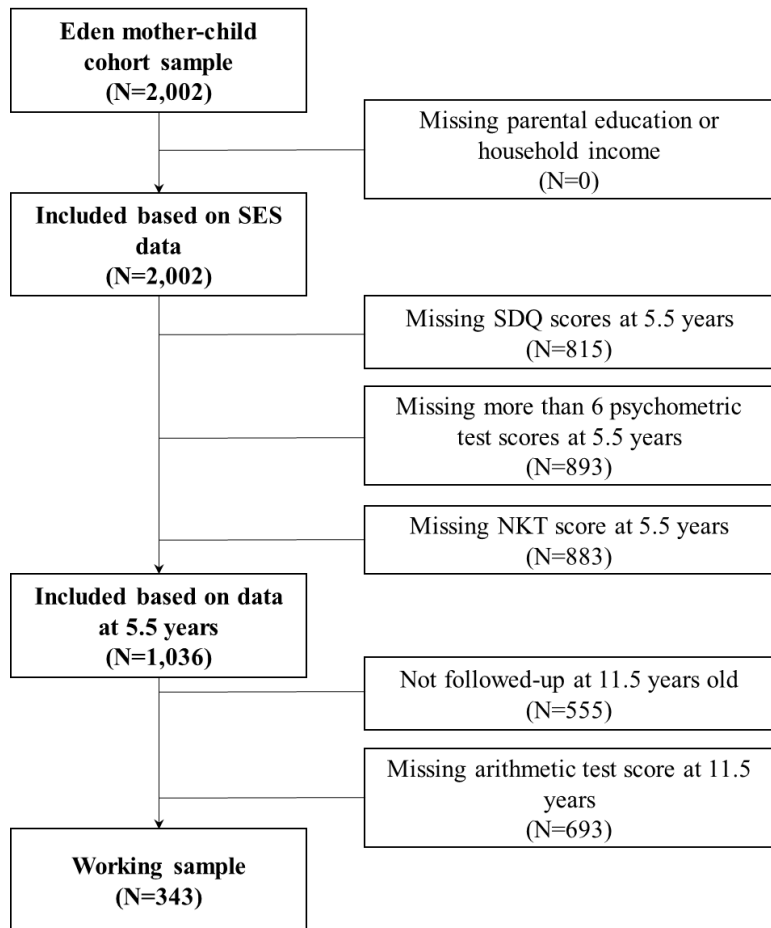


Table A1*Descriptive statistics of included and excluded participants*

Variable	Excluded (present at 5.5 years)			Included (present at 11.5 years)			Difference
	N	Mean	SD	N	Mean	SD	p-value
NKT (out of 18)	776	11.20	3.86	343	12.45	3.42	<.0001
Peg moving task (time in seconds)	780	28.50	5.40	341	27.66	4.70	0.0089
Nonword repetition (NEPSY, out of 46)	780	27.52	8.03	342	29.43	7.75	0.0002
Word-segment recognition (NEPSY, out of 14)	769	10.93	1.91	338	11.00	1.78	0.5384
Sentence repetition (NEPSY, out of 34)	760	14.97	4.06	328	16.47	3.99	<.0001
Design copying (NEPSY, out of 72)	771	50.22	7.69	338	52.83	7.23	<.0001
Information (WPPSI-III, out of 34)	767	24.61	3.01	343	25.75	2.66	<.0001
Vocabulary (WPPSI-III, out of 43)	766	23.20	5.67	342	24.60	5.51	0.0001
Word reasoning (WPPSI-III, out of 28)	766	15.63	4.80	342	17.06	4.28	<.0001
Block design (WPPSI-III, out of 40)	766	27.55	3.85	341	29.12	3.69	<.0001
Matrix reasoning (WPPSI-III, out of 29)	766	14.96	3.78	343	16.28	3.97	<.0001
Picture concepts (WPPSI-III, out of 28)	765	13.96	3.99	342	14.92	3.79	0.0002
Coding (WPPSI-III, out of 65)	764	29.20	10.17	339	30.83	9.13	0.0088
Emotional symptoms (SDQ, out of 15)	843	2.10	1.89	343	2.20	1.87	0.4028
Conduct problems (SDQ, out of 15)	842	2.44	2.10	343	2.17	1.89	0.0365
Hyperactivity/inattention (SDQ, out of 15)	841	3.19	2.39	343	2.79	2.38	0.0092
Peer relationship problems (SDQ, out of 15)	842	1.25	1.36	343	1.08	1.22	0.0391
Prosocial behavior (SDQ, out of 15)	843	8.34	1.67	341	8.43	1.69	0.4063
Parental education (years)	1659	13.05	2.35	343	14.25	2.18	<.0001
Household income (k€)	1579	2505.57	1080.23	343	2885.34	914.19	<.0001
Number of siblings at 5.5 years old	1562	1.12	1.01	343	1.23	0.80	0.0357

Chapter 3 - Early predictors of arithmetic skills

Sex (1 = Male)	1560	0.53	0.50	343	0.49	0.50	0.1138
Gestational age (weeks)	1562	39.19	1.79	343	39.36	1.58	0.1030
Birthweight (kg)	1556	3.26	0.52	343	3.35	0.48	<.0001
Mother's age (years)	1659	28.78	5.00	343	29.93	4.70	<.0001
Alcohol consumption during pregnancy (nb drinks/week)	1578	0.52	1.46	343	0.52	1.41	0.9591
Tobacco consumption during pregnancy (1=Yes)	1581	0.30	0.46	343	0.14	0.34	<.0001
Breastfeeding duration (1=more than 3 days)	1549	0.71	0.45	343	0.78	0.42	0.0164
Cognitive stimulation (HOME score, out of 27)	770	17.14	2.35	335	17.42	2.13	0.0478
Maternal depression episode until 5.5 years old (1=Yes)	1405	0.30	0.46	343	0.33	0.47	0.3322
Recruitment centre	1659	1.53	0.50	343	1.47	0.50	0.0552

3.2 Cognitive predictors of multiplication, addition and subtraction

This section is based on the following article:

Guez, A., Pinheiro-Chagas, P., Peyre, H., Ramus, F. & Piazza, M. (*in prep.*). Differential cognitive predictors of multiplication, addition and subtraction skills in 11.5-year-old children.

3.2.1 Abstract

A substantial and converging body of evidence from neuroimaging, behavioral and neuropsychology studies suggests that different arithmetic operations rely on distinct neuro-cognitive processes: while addition and subtraction may rely more on visuospatial reasoning, multiplication would depend more on verbal abilities. In this paper, we tested this hypothesis in a longitudinal study measuring language and visuospatial skills in 358 young preschoolers, and testing their mental calculation skills at the beginning of middle school. We found that language skills at 5.5 years old significantly predicted multiplication, but not addition nor subtraction scores at 11.5 years old. Conversely, early visuospatial skills predicted addition and subtraction, but not multiplication scores. These results provide strong support for the existence of a double dissociation in mental arithmetic operations, and demonstrate the existence of long-lasting links between language/visuospatial skills and specific calculation abilities. Beyond a mechanistic understanding of mental arithmetic, our results provide potential venues for improving mathematical learning in young children.

Keywords: language, visuospatial, mental calculation, longitudinal

3.2.2 Introduction

Behavioral and neuroimaging evidence suggest that different types of arithmetical operations rely on partially different cognitive processes and partially segregated brain circuits (Dehaene et al., 2003). On the one hand, subtraction and addition operations rely more on nonverbal quantitative representation of numbers, underlain by visuospatial abilities. At the neural level, this is associated with an increased activity in mid and posterior parietal activation bilaterally, typically associated with quantity and spatial processing, during subtraction and addition, compared to multiplication (Lee, 2000; Prado et al., 2011; Zhou et al., 2007). On the other hand, multiplication tables are typically rote learnt and stored in verbal memory (Verguts & Fias, 2005). This is reflected by the fact that solving mental multiplication problems (compared to subtractions and additions) results in increased activation in regions involved in verbal processing, such as the left angular gyrus, and the inferior frontal and middle temporal gyri of the left hemisphere (Lee, 2000; Prado et al., 2011; Zhou et al., 2007). Interestingly, such dissociation between operations is so powerful that even the mere view of the arithmetical signs (“+”, “-”, or “×”) triggers different responses : when subjects are presented with addition or subtraction, but not multiplication signs, even in the absence of any arithmetical operation to perform, they engage in involuntary shifts of visuospatial attention along the horizontal plane (Li et al., 2018). Moreover, they activate regions of the posterior parietal cortex linked to spatial attention shifts (Mathieu et al., 2018).

These results are corroborated by behavioral evidence suggesting that addition and subtraction are more closely related to the visuospatial sketchpad, while multiplication is more strongly associated with the phonological loop. For example, Lee & Kang, (2002) found that the simultaneous performance of a visuospatial memory task affected subtractions, but not multiplications; while the simultaneous performance of a phonological memory task affected multiplications, but not subtractions. In the same line, visuospatial working memory tasks appear to predict more variance in addition and subtraction compared to multiplication in the early years of primary school children (van der Ven et al., 2013). On the contrary, temporal and frontal cortex activation observed during a phonological processing task have been found to predict progression in multiplication between 10 and 12 years old, but not in subtraction (Suárez-Pellicioni et al., 2019); and phonological awareness seems to correlate with multiplication and retrieval problems, but not with procedural problems involving additions and subtractions (De Smedt et al., 2010).

In neuropsychology, double dissociations have been reported. On one side, patients can be selectively impaired in subtraction and quantity manipulation but completely spared in multiplication fact retrieval (Dehaene & Cohen, 1997; van Harskamp & Cipolotti, 2001). On the other side, some patients present a selective impairment in multiplication fact retrieval, while concurrently remaining able to solve addition and subtraction problems (Cappelletti et al., 2001; Cohen & Dehaene, 2000; Sandrini et al., 2003; van Harskamp & Cipolotti, 2001). Finally, data from individuals with developmental disorders also appears in line with this model: children and adults with impairments in phonological processing (dyslexic individuals) show marked difficulties in multiplication fact retrieval but no impairment in subtractions (Boets & De Smedt, 2010; De Smedt & Boets, 2010; Simmons & Singleton, 2008). Moreover, the number of trials correctly solved by means of a retrieval procedure has been found to be positively correlated with their degree of phonological awareness (De Smedt & Boets, 2010).

Taken together, these results provide a large body of evidence pointing towards a differentiated effect of visuospatial and language abilities on mental calculation, with visuospatial abilities supporting addition and subtraction, and language supporting multiplication. However, very little is known about the origin of these dissociations. From a developmental standpoint, this suggests that better visuospatial abilities would enhance the acquisition of addition and subtraction, while language abilities would foster the acquisition of multiplication skills. Surprisingly, no study so far has directly tested this hypothesis. Additionally, most of the previous findings are limited by their small sample size ($N \leq 50$; with the notable exception of van der Ven et al., 2013) and cross-sectional design – longitudinal studies examining the association between early visuospatial/language abilities and later calculation skills are scarce (but see Suárez-Pellicioni et al., 2019 for the effect of previous phonological processing abilities on later multiplication and subtraction). Yet, longitudinal evidence is a key step towards establishing causality, by informing and directing future intervention studies. Previous longitudinal studies indicated that both early visuospatial (Yang et al., 2019; Zhang et al., 2014, 2017) and language abilities (Durand et al., 2005; Zhang et al., 2017) play an important role in the acquisition of arithmetic skills, but whether the two sets of abilities have a differentiated influence on the different types of arithmetic operations has not been investigated.

In this paper, we examined whether the association between arithmetical computation and visuospatial skills, and that between multiplication fact retrieval and

language, are also reflected in developmental, longitudinal data. Our study relies on large sample size (N=358) and a long-range longitudinal approach, where visuospatial and language skills were measured in preschool (T1), way before kids acquire mental arithmetic (measured at T2, in middle-school). In line with previous findings, we hypothesized that there is a specificity in the longitudinal predictors of multiplication versus addition and subtraction, such that: a) early visuospatial skills predict subtraction and addition scores more than multiplication; and b) early language skills predict mental multiplication more than addition and subtraction.

3.2.3 Method

Sample

The data analyzed come from the Eden mother-child cohort (Heude et al., 2016). The initial recruitment sample consisted in 2,002 pregnant women seen during a prenatal visit at the departments of Obstetrics and Gynecology of the French University Hospitals of Nancy and Poitiers before their twenty-fourth week of amenorrhea, who agreed to participate and matched the inclusion criteria. Women with a personal history of diabetes, twin pregnancy, intention to deliver outside the university hospital or to move out of the study region within the following 3 years, and who could not speak French, were excluded from the sample. The participation rate among eligible women was 53%. Enrolment started in February 2003 in Poitiers and in September 2003 in Nancy and lasted for 27 months in each center. 1,907 women out of 2,002 were still in the cohort at delivery. Detailed data on children's environment and cognitive development were regularly collected from birth to 11.5 years old (age at the last wave: Mean=11.56, SD=0.51), with progressive attrition (the numbers of participants at ages 1, 2, 3, 5.5, 8 and 11.5 years of the child were, respectively, 1,717, 1,611, 1,527, 1,255, 883, and 538). At the age of 11.5 years old, 358 students completed the mental calculation test and at least half of psychometric tests at 5.5 years old, which were the conditions for inclusion in our study. Characteristics of this working sample are reported in **Table 1**.

Table 1*Child characteristics*

Variable	N	Mean	S.D.	Median	Min	Max
Parental education (years)	358	14.23	2.17	14.00	10.00	17.00
Age at last test (years)	347	11.57	0.52	11.52	10.49	13.28
Sex (Male, %)	358	48.32				
Schooled in grade 5 (%)	358	20.39				
Schooled in grade 6 (%)	358	58.94				
Schooled in grade 7 (%)	358	19.83				

The study was approved by the Ethical Research Committee (Comité consultatif de protection des personnes dans la recherche biomédicale) of Bicêtre Hospital and by the Data Protection Authority (Commission Nationale de l'Informatique et des Libertés). Informed written consents were obtained from parents for themselves at the time of enrollment and for the newborn after delivery.

Measures

Mental calculation test at 11.5 years old. At the last data collection session, children were administered an online mental calculation test, taken at home on the family computer, as part of larger test battery. It consisted of 24 mental calculation problems of increasing difficulty, including 8 additions, 8 subtractions (both with 1- to 3-digit operands), and 8 single digit multiplications. Calculations at the left hand side of the equal sign were presented in written form (not spoken), and children had to solve them mentally and then type in their answer on the keyboard (in digits) within 10 seconds. We computed an Addition score, a Subtraction score, and a Multiplication score as the sum of correct answers for each category. Answers that were not delivered within 10 seconds were considered wrong. Due to the fact that children were not in the same school grade when tested (ranging from grade 5 to grade 7, see **Table 1**), which was related to their performance at the test, we adjusted the final score for school grade (taking the residuals from the regression of the raw score on school grade). Example of items:

$$\begin{array}{lll}
 8 + 5 = & 7 + 62 = & 245 + 73 = \\
 9 - 4 = & 44 - 8 = & 157 - 13 = \\
 3 \times 6 = & 4 \times 7 = & 9 \times 8 =
 \end{array}$$

Cognitive tests and questionnaires at 5.5 years old. At 5.5 years old, children's cognitive abilities were assessed at home by a trained psychologist, by means of a range of psychometric tests. All psychometric test scores were adjusted on the child's age.

Specific cognitive abilities. Subtests from the WISC-III (Wechsler, 1967, 2004) and NEPSY (Kemp et al., 2001; Korkman et al., 2003) batteries were administered, as well as the Peg-moving task (Nunes et al., 2008).

- *Non-word repetition (NEPSY):* This test is scored as the number of syllables repeated correctly (out of 46 syllables in 13 non-words). It taps phonological processing (encoding and decoding) and verbal short-term memory.

- *Word segment recognition (NEPSY):* Children had to identify pictures that represent words (named in the first place by the psychologist) formed from orally presented word segments. This test taps phonological processing (phonemic awareness).

- *Sentence repetition (NEPSY):* This test is scored as the number of sentences (out of 17) repeated correctly. This test is designed to measure syntactic skills and verbal short-term memory.

- *Design-copying task (NEPSY):* Children had to copy 18 two-dimensional figures correctly (each item was rated from 0 to 4). This test taps visual perception and organization and visual-motor coordination.

- *Information (WPPSI-III):* Children had to correctly answer (verbally or by pointing) 34 questions that address a broad range of general knowledge topics (34 items). This test measures language comprehension, conceptual knowledge and verbal expressive ability.

- *Vocabulary (WPPSI-III):* Children had to correctly define 25 words. This test is designed to measure receptive vocabulary, conceptual knowledge and verbal expressive ability.

- *Word reasoning (WPPSI-III):* Children had to correctly identify a concept from a series of clues (28 items). This test taps language comprehension, conceptual knowledge and general reasoning ability.

- *Block design (WPPSI-III):* Children had to correctly recreate two-dimensional designs using blocks (20 items). This test is designed to measure nonverbal concept formation, visual perception and organization and visual-motor coordination.

- *Matrix reasoning (WPPSI-III):* Children had to correctly complete 29 matrices correctly completed (29 items). This test taps nonverbal concept formation and visual perception and organization.

- *Picture concepts (WPPSI-III)*: Children had to correctly select 2 or 3 pictures with common characteristics (28 items). This test is designed to measure abstract categorical reasoning ability.

Analyses

Data processing was performed with the software SAS 9.4, and statistical analyses were performed with the software Mplus 8. Two latent factors for early language and visuospatial skills respectively were constructed from the 10 psychometric tests, based on results from factor analyses in a previous study (Chapter 3, Part 3.1), with the language latent factor loading on non-word repetition, word-segment recognition, sentence repetition, information, vocabulary, word reasoning and picture concepts, and the visuospatial latent factor loading on matrices, block design, design copying, the peg-moving task, coding, and picture concepts (see **Table 3**). We ran structural equation models (SEM) with these two latent factors as concurrent predictors and the three mental calculation scores as outcomes (see **Figure 1**). Language and visuospatial skills at 5.5 years old were allowed to covary, as well as the addition, subtraction and multiplication scores at 11.5 years old. In the first baseline model (Model a), we only controlled for the recruitment center, by regressing latent cognitive abilities and mental calculation scores on the recruitment center within the model. In a second model (Model b), we controlled in addition for parental education and sex in the same way.

The models' goodness of fit was examined using the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean squared error of approximation (RMSEA). CFI and TLI values greater than 0.95 and values of RMSEA less than 0.06 were used as cut-offs. We only interpreted standardized coefficients in order to compare the effect sizes of the different predictors. The significance of coefficients was assessed with two-sided t-tests. We accounted for multiple comparison with the Benjamini-Hochberg False Discovery Rate (FDR), setting the total number of tests equal to the number of regression coefficients of interest estimated (9) and the false discovery threshold q to 0.05. We excluded from the analysis children who had missing data on at least one of the mental calculation score, and on more than 6 psychometric tests (remaining $N=358$). The remaining missing data was treated using full information maximum likelihood (FIML) estimation.

3.2.4 Results

Descriptive statistics

Table 1 presents the characteristics of the distribution of the three mental calculation raw scores at 11.5 years old as well as the twelve psychometric tests at 5.5 years old. **Table 2** presents the correlations between the three mental calculations scores.

Table 1

Descriptive statistics of tests at 11.5 and 5.5 years old

Variable	N	Mean	S.D.	Median	Min	Max
<i>Tests at 11.5 years old</i>						
Additions (out of 8)	358	5.02	1.53	5.00	1.00	8.00
Subtraction (out of 8)	358	4.68	1.43	5.00	1.00	8.00
Multiplication (out of 8)	358	6.63	1.45	7.00	1.00	8.00
<i>Tests at 5.5 years old</i>						
Nonword repetition (NEPSY)	356	29.31	7.72	30.00	5.00	45.00
Word-segment recognition (NEPSY)	352	10.99	1.78	11.00	4.20	14.00
Sentence repetition (NEPSY)	341	16.43	3.96	16.00	7.00	28.00
Design copying (NEPSY)	351	52.85	7.19	52.80	36.00	69.60
Information (WPPSI-III)	358	25.66	2.70	26.00	17.00	31.00
Vocabulary (WPPSI-III)	357	24.42	5.55	24.00	8.00	40.00
Word reasoning (WPPSI-III)	357	17.01	4.29	18.00	0.00	27.00
Block design (WPPSI-III)	356	29.08	3.65	30.00	20.00	40.00
Matrix reasoning (WPPSI-III)	358	16.21	3.96	16.00	6.00	28.00
Picture concepts (WPPSI-III)	357	14.91	3.79	15.00	5.00	24.00

Table 2*Correlations between the three mental calculation scores at 11.5 years old*

	Addition	Subtraction	Multiplication
Addition	1.00	0.52	0.38
Subtraction		1.00	0.39
Multiplication			1.00

Note: Adjusted scores (on the child's school grade)

Early cognitive predictors of mental calculation skills at 11.5

Table 3 presents the factor structure of the 10 psychometric tasks at 5.5 years old. Results from the ensuing structural equation models predicting the three mental calculation tasks at 11.5 years old with latent cognitive skills at 5.5 years old are presented in **Table 4** and illustrated in **Figure 1**. Latent visuospatial skills were a significant predictor of addition and subtraction skills, but not of multiplication skills (Model a). Thus, an increase in visuospatial skills at 5.5 years old by 1 SD predicted an increase of 0.3 SD in addition and subtraction scores at 11.5 years old, but not in multiplication scores. On the contrary, latent language skills at 5.5 was a significant predictor of multiplication skills at 11.5, but not of addition nor subtraction skills. The same trends were observed when controlling for parental education and sex, with slightly smaller coefficients (Model b).

Table 3*Confirmatory factor analysis of the 10 psychometric test scores at 5.5 years old*

Measured variables	Latent factors			
	Language skills		Visual-spatial skills	
	Estimate	S.E.	Estimate	S.E.
Non-word repetition	0.55	0.02		
Sentence repetition	0.71	0.02		
Word-segment recognition	0.33	0.03		
Information	0.82	0.01		
Vocabulary	0.72	0.02		
Word Reasoning	0.86	0.01		
Picture concepts	0.20	0.04	0.34	0.05
Design copying			0.55	0.03
Block design			0.73	0.03
Matrix reasoning			0.59	0.03
Correlations between latent factors				
Language skills	1.00		0.60	
Visual-spatial skills			1.00	

Note: CFI = 0.988; TLI = 0.983; RMSEA = 0.034 [0.024;0.045]. Nonword repetition and sentence repetition were allowed to covary ($r=0.38$).

Table 4

Results from the structural equation model with the three mental calculation tasks at 11.5 years old predicted by latent cognitive skills at 5.5 years old

Latent predictors at 5.5 years old	Addition		Subtraction		Multiplication	
	β (S.D.)	p-value	β (S.D.)	p-value	β (S.D.)	p-value
Model a						
Language skills	-0.06 (0.07)	0.386	0.01 (0.07)	0.861	0.21* (0.07)	0.002
Visual-spatial skills	0.32* (0.08)	<0.0001	0.27* (0.08)	0.001	0.10 (0.08)	0.204
Model b						
Language skills	-0.06 (0.07)	0.338	0.00 (0.07)	0.999	0.19* (0.07)	0.005
Visual-spatial skills	0.28* (0.07)	<0.0001	0.23* (0.07)	0.002	0.06* (0.08)	0.392

Note: Mental calculation scores and cognitive predictors are solely adjusted on recruitment center in Model a; and on recruitment center, parental education and sex in Model b. * denotes significance after FDR correction with $q=0.05$ and the number of tests equal to 9. Standardized coefficients are reported. The addition, subtraction and multiplication scores were simultaneously entered as dependent variables in the model and were allowed to covary; and the latent cognitive predictors were allowed to covary. Model a: CFI = 0.986; TLI = 0.978; RMSEA = 0.029 [0.000;0.047]. Model b: CFI = 0.981; TLI = 0.961; RMSEA = 0.036 [0.017;0.052].

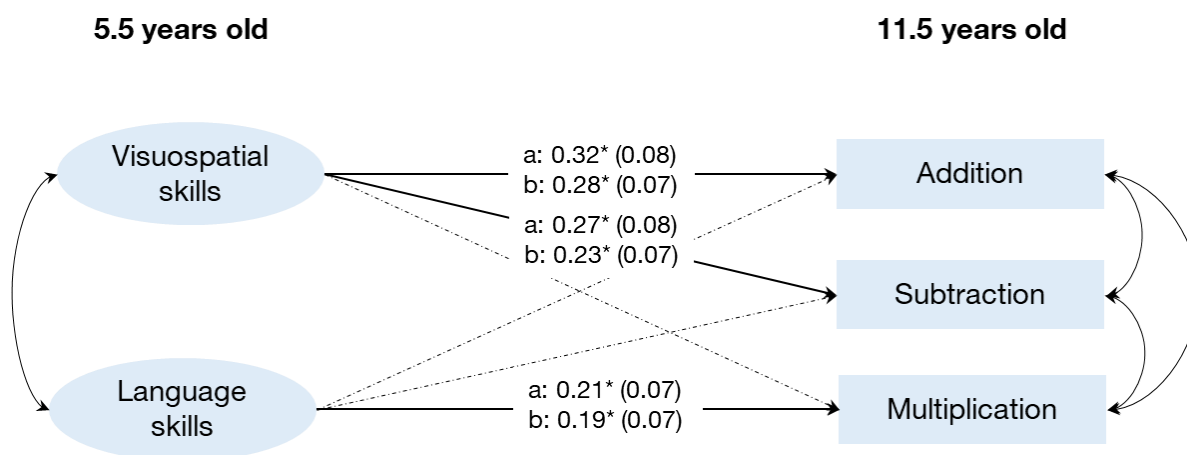


Figure 1: Results from structural equation models with mental calculation tasks as outcomes and latent cognitive skills as predictors. In Model a, mental calculation scores and cognitive predictors are adjusted on recruitment center; in Model b they are adjusted on recruitment center, parental education and sex in Model b. * denotes significance after FDR correction with $q=0.05$ and the number of tests equal to 9. Standardized coefficients are reported.

3.2.5 Discussion

A substantial and converging body of evidence from neuroimaging, behavioral and neuropsychology studies suggest that all mental calculation tasks are not supported by the same neuro-cognitive processes. While addition and subtraction seem to rely more on visuospatial functions, multiplication depends more on verbal abilities. It is thought that this double dissociation, mainly observed through correlational approaches, is caused by the fact that the two types of arithmetical operations are formally taught, at least in Western school, in very different ways: while additions and subtractions are taught and stored through visuospatial supports, such as counting or number lines, multiplication tables are typically learnt by rote in the form of verbal associations. This predicts that the early inter-individual variations across children in visuospatial and language skills before they enter formal schooling should be predictive of later proficiency in solving arithmetical operations, as learnt at school.

In order to test this prediction, we implemented a longitudinal paradigm where we measured language and visuospatial skills in 358 young preschoolers before they were trained in mental arithmetic, and then tested them at the beginning of middle school to measure their calculation skills. We found that visuospatial skills at 5.5 years old significantly predicted later addition and subtraction scores, but not multiplication scores at 11.5 years old. Conversely, early language skills predicted later multiplication scores, but not addition nor subtraction. Thus, these results provide a strong support for the existence of a double dissociation in mental arithmetic operations (Dehaene et al., 2003). Furthermore, we show that this dissociation not only exists concurrently, but also longitudinally: children with better early visuospatial abilities are more likely to compute additions and subtractions correctly 6 years later, while those with better early language abilities are more likely to retrieve multiplication facts correctly 6 years later. These findings thus considerably refine the current knowledge and implications from longitudinal data that early language and visuospatial skills are important building blocks for the acquisition of arithmetic abilities (Durand et al., 2005; Yang et al., 2019; Zhang et al., 2017). They demonstrate the existence of long-lasting and differentiated links, with a specific directionality, between these functions.

Research in the last decades has casted light on the mechanisms that may be underlying this dissociation. While multiplication facts are typically rote-learnt and solved through the use of a retrieval strategy, the processes lying beneath even elementary

additions and subtractions appear to be more complex. Contrary to the previously established belief that simple arithmetic calculations are also solved primarily by direct fact retrieval (Ashcraft, 1992; LeFevre et al., 1996; Siegler & Shrager, 1984), more recent research has provided evidence that basic additions and subtractions are performed by fast automated procedures relying on a spatially organized mental representation of numbers (Barrouillet & Thevenot, 2013; Uittenhove et al., 2016). In this framework, symbolic and non-symbolic additions and subtractions would thus be solved respectively through rightward and leftward shifts along a mental number line (Knops et al., 2009; McCrink et al., 2007; Pinhas & Fischer, 2008; Pinheiro-Chagas et al., 2017, 2018).

These results need to be interpreted with the familiar limitations associated with the use of cohort studies. First, this study remains correlational, due to the observational nature of our data. While it is assumed that language and visuospatial abilities play a causal role in determining later proficiency in mental arithmetic abilities at school, it is also possible that, in turn, the cultural acquisition of mental arithmetic contributes to improving language and visuospatial abilities, in a form of circular causality that also occurs in other domains (e.g., Hulme et al., 2012; Piazza et al., 2010). Our finding that preschool language and visuospatial abilities are differentially associated with middle-school calculation skills is remarkable in itself and suggest the existence of a causal link; yet, our design is not sufficient to establish causal inference. This result thus calls for a further investigation of the purported causal links between language/visuospatial skills and different calculation skills through intervention studies. A second limitation concerns the representativeness of our sample. Indeed, while our sample size is large compared to most studies in the domain, it is worth considering that its external validity may not be complete given the selective attrition between inclusion in the Eden cohort and the 11.5 years old wave. Indeed, children present at 11.5 years old have more highly educated parents and higher cognitive abilities than those present at the 5.5 years wave who could not be followed-up and included in our sample (see Appendix). However, this limitation applies to the vast majority of studies in the domain, where representativeness of the tested sample relative to the whole population is often very hard to achieve.

In spite of these limitations, our study provides strong evidence, from a large longitudinal sample, that visuospatial and language abilities measured prior formal mathematics instruction differentially predict addition, subtraction, and multiplication. These results suggest that training early visuospatial skills may enhance later addition and subtraction abilities but have no influence on multiplication abilities; conversely training

Chapter 3 - Early predictors of arithmetic skills

early language skills may benefit the later acquisition of multiplication skills, but not addition nor subtraction. Our results may also have important implications in designing re-education tools for children with developmental learning disabilities: in children with dyscalculia and concurrent weak visuospatial skills, re-education might be based on extending rote verbal learning of simple multiplication tables to simple additions and subtractions. On the contrary, dyscalculic children with associated language impairments might profit from using visuospatial strategies for memorizing simple multiplication tables.

Chapter 4 - Early predictors of literacy skills

4.1 Cognitive and environmental predictors of reading and spelling

This section is based on the following article:

Guez, A.*, Di Folco, C.*, Peyre, H., & Ramus, F. (*in prep.*). Early cognitive and environmental predictors of literacy skills at 11.5 years old.

4.1.1 Abstract

Learning to read is crucial for the development of individuals and their insertion in society. Reading is acquired throughout schooling, but its foundations are laid long before instruction begins. Understanding which early cognitive, socio-emotional and environmental characteristics underlie reading acquisition is therefore of great importance for actors in the education sphere. Here, we study the role of a large set of such predictors in the EDEN cohort, following 297 children from age 5.5 (assessment of cognitive and socio-emotional factors) to 11.5 (assessment of reading comprehension, reading accuracy, reading speed and spelling). The role of environmental predictors was further investigated in a mediation analysis, allowing to disentangle their direct effects on literacy skills from their indirect effects through early cognitive and socio-emotional variables. Phonological processing and short-term memory at 5.5 years old predicted reading accuracy, while short-term memory and conceptual knowledge predicted spelling. Besides, visuospatial and fine motor abilities did not predict literacy outcomes beyond the role of language. Moreover, hyperactivity/inattention was negatively associated with reading accuracy. Lastly, parental education and cognitive stimulation predicted reading accuracy, and parental education further predicted reading comprehension. The effect of parental education on literacy at 11.5 seemed to be part direct and part mediated by early language skills, while the effect of cognitive stimulation was mostly mediated by early cognitive skills. Our final model respectively explained 41.5%, 30.6%, 29% and 22.7% of the variance of reading

Chapter 4 - Early predictors of literacy skills

comprehension, accuracy, spelling and reading speed. These results shed light on how a wide set of factors drive the development of literacy in children, and the mechanisms of action of the early environment.

Keywords: literacy, early predictors, language abilities, socio-emotional symptoms, home learning environment, parental education

4.1.2 Introduction

Learning to read is a crucial step in the development of individuals, enabling the acquisition of new knowledge and thereby allowing active participation in society (Castles et al., 2018) as well as better educational and professional outcomes (Hahn & Truman, 2015). However, more than 20% of teenagers in OECD countries have still not attained a basic proficiency level (OECD, 2016a, 2019). In France, the situation is similar: in 2007, 21% of pupils left primary school with great difficulties in reading, this proportion having increased compared to that of the preceding decade (Rocher, 2008). France is also one of the countries where the gap between students with high- and low- reading skills is the highest, these inequalities drawing heavily upon social backgrounds (OCDE, 2012). Learning to read builds on cognitive abilities developed before formal instruction begins, and the differences then created hardly resolve over time (Landerl & Wimmer, 2008). A better understanding of the mechanisms underlying reading acquisition, including its most precocious predictors, is therefore critical to improve educational practices and ultimately guarantee literacy for all.

Early language predictors of different literacy skills

Language skills form the most important basis of literacy learning; however, different literacy skills have been shown to have different early language predictors (Moll et al., 2014; Storch & Whitehurst, 2002). As such, one must differentiate word recognition, reading comprehension and spelling skills.

Word recognition is the ability to translate printed input into speech sounds and to recognize written words accurately (Durand et al., 2013). It is traditionally assessed by single word or pseudo word reading while measuring speed and accuracy. According to the

dual-route model, word recognition would include both an “orthographic procedure based on lexical units” and a “sublexical phonological procedure based on grapheme–phoneme correspondences” (Coltheart et al., 2001). Hence, there exists a tight relationship between children’s phonological skills and reading ability (Melby-Lervåg et al., 2012; National Early Literacy Panel, 2008). In particular, phonological awareness (manipulating speech sound units) has been shown to be a strong predictor of word recognition (Hjetland et al., 2020; Muter et al., 2004), the strongest being phonemic awareness (Melby-Lervåg et al., 2012). In addition to phonological awareness, other important predictors of word recognition include storing (verbal short-term memory) and accessing speed (measured by rapid naming) of phonological representations, as well as letter knowledge (Hjetland et al., 2020; Melby-Lervåg et al., 2012). When reading accuracy and speed are distinguished, phonological awareness proved a better predictor for accuracy while rapid naming better predicts speed (Moll et al., 2014). Vocabulary may also facilitate reading fluency, although results are mixed (Su et al., 2017; but see (Muter et al., 2004; Piquard-Kipffer & Sprenger-Charolles, 2013).

Reading comprehension is the ability to extract meaning from written words and sentences, generally assessed by having the child answer questions about a text (Durand et al., 2013). As it constitutes the goal of reading, it has been extensively studied. The “simple view of reading” (Gough & Tunmer, 1986; Hoover & Gough, 1990) is a widely-used theoretical framework according to which reading comprehension is the product of two relatively independent skills: word recognition and (oral) language comprehension, both skills being necessary but not sufficient. While their relative importance is still unclear, decoding is supposed to be a bottleneck in early grades, while language comprehension becomes a better predictor in later grades (Gough et al., 1996). As a crucial component of oral comprehension, vocabulary has thus shown consistent influence on later reading comprehension abilities (Durand et al., 2013; Hjetland, 2018; Hjetland et al., 2020; Kendeou et al., 2009; Muter et al., 2004; NICHD Early Child Care Research Network, 2005; Ouellette, 2006; Protopapas et al., 2012; Senechal, 2006; Storch & Whitehurst, 2002; Su et al., 2017). Other language skills required for language comprehension have been shown to contribute to reading comprehension beyond vocabulary, such as grammar skills (Durand et al., 2013; Hjetland, 2018; Hjetland et al., 2020; Lehl et al., 2020; Muter et al., 2004; NICHD Early Child Care Research Network, 2005; Storch & Whitehurst, 2002; Su

Chapter 4 - Early predictors of literacy skills

et al., 2017) and conceptual knowledge (Hjetland et al., 2020; National Early Literacy Panel, 2008; Storch & Whitehurst, 2002).

Lastly, spelling is an important and complementary component of literacy, which shares a common basis with reading, relying on letter knowledge (Caravolas et al., 2001; Lervåg & Hulme, 2010) and memory for specific words' spelling. It also hinges on phonological abilities, which are here used to transcode phoneme to grapheme, rather than grapheme to phoneme as in reading – both directions having partly different mappings (Ehri, 2000). Thus, spelling is predicted by phoneme awareness (Landerl & Wimmer, 2008; Lervåg & Hulme, 2010), but also by rapid naming (Landerl & Wimmer, 2008; Lervåg & Hulme, 2010; Moll et al., 2014; Savage et al., 2008) and phonological short-term memory (Lervåg & Hulme, 2010; Moll et al., 2014, but see Plaza & Cohen, 2007).

Non-verbal predictors

Apart from language skills, non-verbal cognitive and socio-emotional skills as well as environmental and individual factors appear to have influences on literacy, although their importance is lesser compared to verbal predictors (Elbro & Scarborough, 2004).

Cognitive predictors

Visuospatial abilities are needed for letter identification and segmentation of written letters into phonemes. Deficits in these skills, in particular in visual attention, were indeed observed in children with or at risk of developmental dyslexia (Bosse et al., 2007; Facoetti et al., 2010; Gori et al., 2016; Vidyasagar & Pammer, 2010). However, it is possible that such impairments could be a consequence rather than a cause of reading disorders (Ramus, 2003), and hence not be an early predictor. Though the question has received slightly less attention in non-pathological populations, the National Early Literacy Panel (2008) reported low univariate correlations between reading outcomes and visuospatial skills (around 0.2); Bosse & Valdois (2009) found that visual attention was linked to reading skills even when controlling for phonological abilities; and Valdois et al. (2019) further showed that kindergarten visual attention could predict later reading skills. The role of visual abilities on spelling has also specifically been studied to understand how orthographic learning occurs: Bosse et al. (2015) and Plaza & Cohen (2007) found that

visual attention explains orthographic learning beyond phonological skills (but see Caravolas et al., 2001, for an absence of effect of visual working memory).

A second set of cognitive abilities that could influence literacy acquisition is fine motor skills. Fine motor abilities in kindergarten have indeed been linked with early reading (Cameron et al., 2012; Grissmer et al., 2010; Pitchford et al., 2016). This relationship is poorly understood (Suggate et al., 2018), but could stem from the fact that fine motor skills and reading rely on common cognitive skills, such as attention or memory, or that children with better fine motor skills have more opportunities to engage in learning activities promoting academic success than their peers (Suggate et al., 2019).

Socio-emotional predictors

Socio-emotional skills influence both individual learning and classroom dynamics: they determine the quality of the relation with the teacher (Pianta & Stuhlman, 2004), social inclusion (Newcomb et al., 1993) and participation in collaborative learning activities (Ladd et al., 1999). Two dimensions in socio-emotional disorders are usually distinguished: in externalizing problem behaviors, negative emotions are directed toward others, resulting in conduct disorders or attention deficit and hyperactivity; in internalizing problem behaviors, these emotions are directed at oneself, causing depression and anxiety (Roeser et al., 1998). Comorbidity between reading disability and socio-emotional disorders is well-established (Carroll et al., 2005; Willcutt & Pennington, 2000) and seems to have both genetic and environmental origins (Couto et al., 2009; Hart et al., 2010). Longitudinal studies are scarce but Breslau et al. (2009) linked early socio-emotional problems to lower academic achievement. The effect of externalizing disorders – such as inattention-hyperactivity and conduct disorders – on reading achievement was shown to be at least partly mediated by inattention (Breslau et al., 2009; Carroll et al., 2005; Duncan et al., 2007; Rapport et al., 1999; Willcutt & Pennington, 2000); whether it is also the case for internalizing disorders is less consensual (Carroll et al., 2005; Willcutt & Pennington, 2000).

Environmental/individual predictors

Beyond cognitive and socio-emotional factors, a set of environmental and individual variables influence literacy learning, first of which socio-economic status (SES) (Sirin, 2005), with children from higher socio-economic families performing better than others. Closely linked to SES, the cognitive stimulation a child receives at home also

predicts later reading achievement, even when SES is controlled for (Bradley et al., 2001; Melhuish et al., 2008; Sammons et al., 2015) Apart from SES and cognitive stimulation, the role of other predictors has been investigated – though less intensively. Observational studies have shown that breastfeeding is suspected to have a positive impact on a wide range of cognitive abilities, such as general intelligence (Horta et al., 2015; see Mortensen, 2002 for effects in adulthood) or language skills (Heikkilä et al., 2014; Whitehouse et al., 2011). Moreover, a randomized controlled trial has linked breastfeeding promotion with better achievement at 6 years of age (Kramer, 2008). This effect could be attributable to better mother-infant interaction through breastfeeding (Britton et al., 2006), reduced likelihood to develop infectious disease (Ip et al., 2007), or presence of nutrients in the milk, such as fatty acids, that fosters brain development (McCann & Ames, 2005). Higher number of siblings has been linked to lower pre-reading (Sammons et al., 2004) and verbal abilities (Havron et al., 2019), possibly due to competition for parent's attention. Moreover, higher mother age (Greenberg et al., 1999), female sex (Lynn & Mikk, 2009; Reilly et al., 2018) and higher birth weight (Sammons et al., 2004) positively predict reading achievement, while early maternal depression (Sohr-Preston & Scaramella, 2006) negatively predicts it. In longitudinal studies, prenatal factors, such as alcohol or cigarette use during pregnancy have also been associated to lower reading scores at the end of primary school and middle school (Fried et al., 1997; Goldschmidt et al., 2004; Streissguth et al., 1994) According to reviews, prenatal tobacco exposition could have effects on global academic achievement and intellectual abilities (Clifford et al., 2012), although this association is markedly reduced when maternal covariates are controlled for (Batty et al., 2006). Prenatal alcohol exposition is associated with more risks of socio-emotional issues, attention, memory and learning issues (Jacobson & Jacobson, 2002).

As these individual and environmental factors influence the child's whole cognitive development, their association with reading achievement could be at least partly mediated by early cognitive abilities. Individual and environmental factors can thus be seen as distal factors, located near the beginning of the causal chain, and acting via intermediary, proximal factors – early cognitive and socio-emotional abilities. As such, mediation analysis have consistently shown that the roles of cognitive stimulation (Lehrl et al., 2020; Senechal & LeFevre, 2002; Su et al., 2017) and SES (Su et al., 2017) on later reading comprehension were almost fully mediated by early language skills (but see Senechal, 2006, for a direct role of parent teaching literacy on reading fluency at 11).

Limits of the literature and the present study

Investigating the relative importance of early predictors of literacy is a central research question, but is made complicated as studies include different predictors, outcomes and age ranges (Kendeou et al., 2009). Most studies have focused on a particular category of predictors (i.e., verbal cognitive predictors, socio-emotional abilities), without taking into account the effect of others. An exception is the study by Fluss et al. (2009), who looked at the influence of SES, behavioral and emotional problems, and cognitive skills, on reading ability (measured at the same time) among 1062 elementary school students in Paris. Their results emphasize large differences in SES, as well as influences of phonological awareness and attention deficits. Few studies have investigated the long-term direct and mediated effects of early environmental factors on literacy (Lehrl et al., 2020; Senechal, 2006; Su et al., 2017; Zhang et al., 2013), and none included other environmental characteristics than SES and cognitive stimulation or socio-emotional problems. Furthermore, studies seldom followed children from kindergarten to grades where they should master reading. A number of studies follow children up to 8 years of age (Durand et al., 2013; Landerl & Wimmer, 2008; NICHD Early Child Care Research Network, 2005; Storch & Whitehurst, 2002; Su et al., 2017; Torppa et al., 2016; Zhang et al., 2013), but few kept up to age 11 (Durand et al., 2013; Landerl & Wimmer, 2008; Su et al., 2017).

Besides, the vast majority of studies were led in English-speaking populations (Hjetland et al., 2020). However, this language is known for being one of the most inconsistent both in grapheme-to-phoneme and phoneme-to-grapheme mappings. French has a similarly inconsistent phoneme-to-grapheme mapping, but its grapheme-to-phoneme mapping is twice as consistent as that of English (Ziegler et al., 1996). Moreover, English, like other Germanic languages, has a high syllable complexity compared to Romance languages like French (Seymour et al., 2003). These characteristics affect not only the rate of reading acquisition (Seymour et al., 2003), especially slow in English, but also the relative weight of predictors on reading performance (Florit & Cain, 2011; Ziegler et al., 2010). For example, the impact of phonological awareness is stronger for inconsistent orthographies (Ziegler et al., 2010); hence principally studying English might have led to an overestimation of the role of phonological awareness (Share, 2008). It is thus difficult to generalise results from English studies to the French situation. French longitudinal studies modelling the acquisition of reading most often started follow-up after instruction began (Gentaz et al., 2015; Goigoux et al., 2016; Massonnié et al., 2019). Studies on early

predictors are very scarce, and are limited by their small sample size. For instance, Piquard-Kipffer & Sprenger-Charolles (2013) found that pre-reading, phonemic segmentation, phonological short-term memory and letter-name knowledge of 85 children at 5 significantly predicted their reading fluency at 8, contrary to vocabulary, syllable segmentation and visual short-term memory. Plaza (2001) studied the impact of kindergarten phonological and syntactic abilities on reading fluency in grade 1 (N=35), and Plaza & Cohen (2007) showed the impact of visual attention and phonological processing on grade 1 reading fluency (N=77). Casalis & Louis-Alexandre (2000) investigated the impact of kindergarten phoneme and morphological awareness on grade 2 reading accuracy, speed and comprehension (N=50). Lastly, Senechal (2006) studied the importance of home literacy environment and early letter knowledge, vocabulary and phoneme awareness on reading fluency, comprehension and spelling in grade 1 and 4 (N = 65). They showed that vocabulary was the best predictor of reading comprehension in grade 4, and phoneme awareness best predicted reading fluency and spelling in grade 1.

In the present paper, we studied the relative roles of a wide range of early predictors (language skills, non-verbal cognitive skills, socio-emotional abilities, environmental and individual characteristics) on reading accuracy, reading speed, reading comprehension, and spelling in 297 French children followed from birth to 11.5 years old, using structural equation modelling. Moreover, we assessed the direct and indirect effects of environmental and individual predictors through cognitive and socio-emotional variables. We investigated the following research questions (pre-registered on OSF¹⁸):

RQ 1. Which pre-school language abilities are the best unique predictors of literacy skills at 11.5 years old?

H1. Reading accuracy, reading speed and spelling are best predicted by phonological awareness, phonological working memory and vocabulary.

H2. Reading comprehension is best predicted by syntactic abilities, vocabulary and verbal intelligence.

RQ2. How do preschool general language skills, visual-spatial skills and fine-motor skills differentially predict acquired literacy skills?

H1. Early general language predicts all aspects of literacy at 11.5 years old, more than visual-spatial skills do.

¹⁸ https://osf.io/bew4u/?view_only=cd310a714a664ec7bdfbf800f2c8931

H2. Early fine-motor skills predict spelling more than reading skills (nature of the task).

RQ3. Do certain early socio-emotional difficulties/abilities (emotional symptoms, peer relationship problems, hyperactivity-inattention, conduct problems, and prosocial behaviour) predict later literacy skills, beyond these early cognitive skills?

H1. Early inattention-hyperactivity symptoms have a negative impact on later literacy skills.

RQ4. Which pre- and post-natal individual and environmental factors (parental education, household income, parental cognitive stimulation, number of siblings, mother's age, gestational age, maternal depression, maternal consumption of tobacco and alcohol during pregnancy, breastfeeding, birth weight, and sex) uniquely predict literacy skills at 11.5 years old?

H1. There is a sex difference in all literacy skills in favour of girls.

H2. Parental education positively predicts all aspects of literacy.

H3. A higher number of older siblings negatively predicts all aspects of literacy.

H4. Breastfeeding positively predicts all aspects of literacy.

RQ5. To what extent are the effects of these unique pre- and post-natal predictors mediated through early cognitive and behavioural skills?

H1. The effect of parental education and sex are largely mediated through early language skills, but also partly direct.

H2. The effects of number of siblings and breastfeeding are fully mediated by early language skills.

4.1.3 Method

Sample

We used data from the EDEN mother-child cohort (Heude et al., 2016). Pregnant women seen during a prenatal visit at the departments of Obstetrics and Gynaecology of the French University Hospitals of Nancy and Poitiers before their twenty-fourth week of gestation were invited to participate. Exclusion criteria included a personal history of diabetes, twin pregnancy, intention to deliver outside the university hospital or to move out of the study region within the following 3 years, and inability to speak French. The

participation rate among eligible women was 53%. Enrolment started in February 2003 in Poitiers and in September 2003 in Nancy and lasted for 27 months in each centre and resulted in the inclusion of 2,002 pregnant women (1,907 still in the cohort at delivery). The study was approved by the Ethical Research Committee (Comité consultatif de protection des personnes dans la recherche biomédicale) of Bicêtre Hospital and by the Data Protection Authority (Commission Nationale de l'Informatique et des Libertés). Informed written consents were obtained from parents for themselves at the time of enrolment and for the new-born after delivery. Detailed data on children's environment and cognitive development were regularly collected from birth to 11.5 years old (age at the last wave: Mean = 11.57, SD = 0.50). We included in this study data from the waves at 11.5 years old (literacy scores), 5.5 years old (cognitive and socio-emotional scores), and prior to 5.5 years old (individual and environmental characteristics from the maternal questionnaires). At the age of 11.5 years old, 539 parents completed the questionnaires and 297 had data meeting our inclusion criteria (displayed in **Figure 1**). The socio-demographic and individual characteristics of our working sample are displayed in **Table 1**. Comparison with excluded participants is presented in **Table A1**.

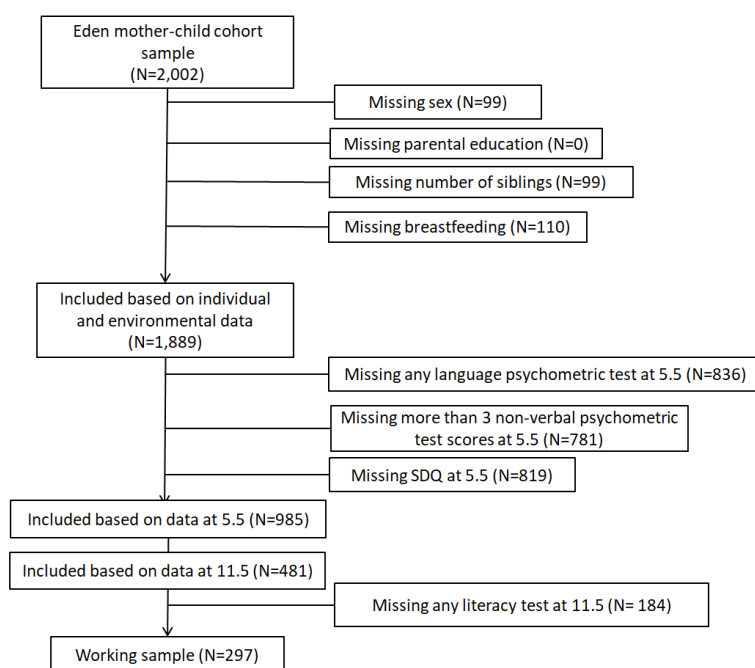


Figure 1: Flowchart of included participants

Measures

Literacy variables at 11.5 years old. Children had to complete 6 literacy tests on the family computer. Due to the fact that they were not all in the same school grade when tested (ranging from grade 5 to grade 7, see **Table 1**), we adjusted the final score for school grade (taking the residuals from the regression of the raw score on school grade).

- *Orthographic choice*: Children had to choose the correct spelling of a word out of three possibilities (13 items). The test is scored as the number of correct answers given. Internal consistency was acceptable (Cronbach's $\alpha = 0.62$)

- *Sentence dictation*: Children were dictated a short text on the phone, part by part, which they had to type correctly on the computer. 10 lexical and 10 grammatical target errors were corrected. An accuracy score was computed. Internal consistency was good (Cronbach's $\alpha = 0.85$).

- *Word dictation*: Children were dictated 50 words (5 blocks: 10 simple regular words, 10 complex regular words, 10 irregular words, 10 2-syllables non-words and 10 3-syllables non-words). A score was computed for each block as the number of correct answers. That is to say, the regular word reading score was computed by adding the two scores for simple and complex regular word dictation and the non-word reading score was computed by adding the two scores for 2- and 3-syllables non-word dictation. Internal consistency was good (Cronbach's $\alpha = 0.81$).

- *Reading comprehension*: Children had to read two short texts and answer 24 multiple choice questions with 3 possible answers (12 questions for each text) to assess their understanding. A score was computed as the number of correct answers. Internal consistency was poor (Cronbach's α respectively equal to 0.40 and 0.58 for each text).

- *Text reading*: Children were asked to read a text in 1 min, on the phone. The number of words read and the number of errors were measured.

- *Word reading*: Children had to read 60 words (3 blocks: 20 words, 20 irregular words, and 20 non-words). For each block, a score was computed as the number of correct answers, and the response time was measured.

Individual items were not available for the reading tests at 11.5, which were administered over the phone (text reading and word reading); therefore, we could not compute reliability estimates for these tests.

Table 1*Socio-demographic and individual characteristics of included participants*

Variable	N	Mean	SD	Min	Median	Max	Kurtosis	Skew
Mother age (years)	297	29.87	4.72	17.00	29.00	42.00	2.80	0.31
Number of older siblings	297	0.70	0.87	0.00	1.00	6.00	9.08	1.86
Parental education (years)	297	14.35	2.15	10.00	14.50	17.00	1.85	-0.26
Household monthly income (k€)	297	2936.48	907.38	975.00	2900.00	5066.70	2.81	0.44
Breastfeeding duration (1=more than 3 days)	297	0.78	0.41	0.00	1.00	1.00	2.92	-1.38
Cognitive stimulation (HOME score, out of 27)	290	17.42	2.12	11.00	18.00	21.00	2.97	-0.57
Maternal depression episode before 5.5 years old (1=yes)	297	0.33	0.47	0.00	0.00	1.00	1.55	0.74
Alcohol consumption during pregnancy (nb drinks/week)	297	0.54	1.47	0.00	0.00	17.00	78.21	7.75
Tobacco consumption during pregnancy (1=yes)	297	0.13	0.34	0.00	0.00	1.00	5.77	2.18
Sex (1=male)	297	0.48	0.50	0.00	0.00	1.00	1.01	0.09
Gestational age (weeks)	297	39.31	1.60	31.00	40.00	42.00	6.89	-1.46
Birth weight (kg)	297	3.33	0.46	1.74	3.33	5.26	4.64	-0.16
Centre (1 = Poitiers, 2 = Nancy)	297	1.47	0.50	1.00	1.00	2.00	1.02	0.13
Schooled in grade 5 (1=Yes)	297	0.20	0.40	0.00	0.00	1.00	3.36	1.54
Schooled in grade 6 (1=Yes)	297	0.59	0.49	0.00	1.00	1.00	1.12	-0.35
Schooled in grade 7 (1=Yes)	297	0.22	0.41	0.00	0.00	1.00	2.85	1.36
Age at the last wave (years)	289	11.55	0.51	10.50	11.50	13.20	3.00	0.48

Cognitive and socio-emotional variables at 5.5 years old. At 5.5 years old, children's cognitive abilities were assessed at home by a trained psychologist, by means of a range of psychometric tests. All psychometric test scores were adjusted for the child's age and birthplace. In addition, mothers filled in a questionnaire on children's mental health.

Language tests. Subtests from the WISC-III (Wechsler, 1967, 2004) and NEPSY (Kemp et al., 2001; Korkman et al., 2003) batteries were administered.

- *Non-word repetition (NEPSY)*: This test is scored as the number of syllables repeated correctly (out of 46 syllables in 13 non-words). It taps phonological processing (encoding and decoding) and verbal short-term memory. Internal consistency was good (Cronbach's $\alpha = 0.80$).

- *Word segment recognition (NEPSY)*: Children had to identify 14 pictures that represent words (named in the first place by the psychologist) formed from orally presented word segments. This test taps phonological processing (phonemic awareness). Internal consistency was poor (Cronbach's $\alpha = 0.43$).

- *Sentence repetition (NEPSY)*: This test is scored as the number of sentences (out of 17) repeated correctly. This test is designed to measure syntactic abilities and verbal short-term memory. Internal consistency was good (Cronbach's $\alpha = 0.77$)

- *Design-copying task (NEPSY)*: Children had to copy 18 two-dimensional figures correctly (each item was rated from 0 to 4). This test taps visual perception and organization and visual-motor coordination. Internal consistency was good (Cronbach's $\alpha = 0.74$).

- *Information (WPPSI-III)*: Children had to correctly answer (verbally or by pointing) 34 questions that address a broad range of general knowledge topics (34 items). This test measures language comprehension, conceptual knowledge and verbal expressive ability.

- *Vocabulary (WPPSI-III)*: Children had to correctly define 25 words. This test is designed to measure receptive vocabulary, conceptual knowledge and verbal expressive ability.

- *Word reasoning (WPPSI-III)*: Children had to correctly identify a concept from a series of clues (28 items). This test taps language comprehension, conceptual knowledge and general reasoning ability.

Nonverbal tests. Subtests from the WISC-III (Wechsler, 1967, 2004) were administered, as well as the Peg-moving task (Nunes et al., 2008).

- *Peg-moving task*: Children had to move five pegs, one by one, in a forward motion pattern, beginning with the peg at the side of each hand. The task started with the preferred hand and the subject had to perform three complete trials with each hand. This test is designed to measure visual-motor coordination. It is scored as the time taken by the child to move the pegs, in seconds. Internal consistency was good (Cronbach's $\alpha = 0.81$).

- *Block design (WPPSI-III)*: Children had to correctly recreate two-dimensional designs using blocks (20 items). This test is designed to measure nonverbal concept formation, visual perception and organization and visual-motor coordination.

- *Matrix reasoning (WPPSI-III)*: Children had to correctly complete 29 matrices correctly completed (29 items). This test taps nonverbal concept formation and visual perception and organization.

- *Picture concepts (WPPSI-III)*: Children had to correctly select 2 or 3 pictures with common characteristics (28 items). This test is designed to measure abstract categorical reasoning ability.

- *Coding (WPPSI-III)*: In limited time (120 seconds), children had to copy symbols, correctly pairing 2 symbols together (59 items). This test taps processing speed, visual-motor coordination and visual working memory.

The manual of WPPSI-III reports evidence of high subtest reliability (0.83 to 0.95), internal consistency, test-retest stability, and validity for all subtests (Wechsler, 1967)¹⁹.

Socio-emotional variables. Mothers answered the Strength and Difficulties Questionnaire (SDQ) (Goodman, 1997, 2001; Shojaei et al., 2009) when the child was 5.5 years old. The SDQ is a screening questionnaire for behavioural problems in children, which includes questions about 25 positive and negative attributes (each item being rated by the mother using a 3-point Likert scale: 1 = “Not true”, 2 = “Somewhat true”, 3 = “Certainly true”), equally divided into five subscales:

- *Emotional symptoms*: e.g. “Many worries, often seems worried”, “Nervous or clingy in new situations, easily loses confidence”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.61$).

- *Conduct problems*: e.g. “Often has temper tantrums or hot tempers”, “Often fights with other children or bullies them”. Internal consistency was good (Cronbach’s $\alpha = 0.74$).

- *Hyperactivity/inattention*: e.g. “Restless, overactive, cannot stay still for long”, “Easily distracted, concentration wanders”. Internal consistency was good (Cronbach’s $\alpha = 0.76$).

- *Peer relationship problems*: e.g. “Rather solitary, tends to play alone”, “Picked on or bullied by other children”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.50$).

- *Prosocial behavior*: e.g. “Considerate of other people’s feelings”, “Kind to younger children”. Internal consistency was acceptable (Cronbach’s $\alpha = 0.69$).

The reliability estimates in our sample were similar to those found in a representative sample of 1,348 French children aged 6–11 years old (Shojaei et al., 2009).

¹⁹ We did not have access to the data per item for the WPPSI-III, therefore we could not compute reliability estimates for our sample.

Environmental and individual variables.

Individual characteristics and birth factors. Child sex (coded as 0 for female sex and 1 for male sex), gestational age, birth weight, and the maternity hospital in which the mother was recruited and the child was born (Nancy or Poitiers), were reported.

Prenatal factors. Mothers indicated their tobacco and alcohol consumption during pregnancy. Tobacco consumption was dummy coded (1 when the mother smoked and 0 otherwise) and alcohol consumption was continuous (number of glasses per week).

Family characteristics and post-natal factors. Mothers reported their age, their own and the father's level of education (a continuous variable equal to the average of both parents' years of education was then constructed), their household income, the child's number of siblings, and breastfeeding duration (a dummy variable equal to 1 if the child was breastfed at least 3 days, 0 otherwise, was constructed). Parental cognitive stimulation of the child at home was assessed by a psychologist using three subscales of the Home Observation for the Measurement of the Environment inventory: language stimulation, academic stimulation, and variety of experimentations (27 items) (Bradley & Caldwell, 1984; Frankenburg & Coons, 1986). A HOME score was computed as the mean of answers to all items. Lastly, mothers answered the Centre for Epidemiologic Studies – Depression (CES-D) questionnaire (Radloff, 1977) (20 items) when the child was aged 4 months, 8 months, 1 year, 3 years and 5.5 years. Scores range from 0 to 60, with high scores indicating greater depressive symptoms. A cut-off score (16 or greater) is normally used in order to identify mothers at risk for clinical depression. A dummy-coded variable indicating whether the mother had at least one depressive episode between child age 4 months and 5 years, was constructed.

Analyses

Our analysis plan was preregistered on OSF. The slight differences between the current analysis and the preregistered analysis are indicated in the main text below. Data processing and analysis were performed with R 3.6.2; Structural Equation Models (SEM) were computed with the R lavaan package 0.6-5.

Models. We first conducted a confirmatory factor analysis with all literacy skills at 11.5 years old to obtain four latent literacy skills (reading speed, reading accuracy, reading

comprehension, and spelling), which served as our outcomes. We then ran the following five structural equation models, answering each of our research questions:

- *Model A*: We constructed a SEM with all latent literacy variables as outcomes and all language test scores at 5.5 years old as predictors. This model enabled us to disentangle the unique contributions of specific language skills to different literacy skills at 11.5.

- *Model B*: We constructed a SEM with all latent literacy variables as outcomes and three latent cognitive abilities (language, visuospatial and fine motor skills) at 5.5 years old as predictors. The latent cognitive abilities measurement model was based on results from factor analyses of all psychometric test scores at 5.5 years old conducted in a previous study (Chapter 3, Part 3.1). This model allowed us to investigate the specific effect of early language, fine motor and visuospatial skills on literacy skills at 11.5.

- *Model C*: We added to Model B the five SDQ subscales at 5.5 years old as predictors of literacy skills at 11.5 years old. This model assessed the effects of inattention-hyperactivity, emotional symptoms, peer relationship issues, conduct problems and prosocial behaviour at 5.5 on literacy skills at 11.5, beyond the effects of early cognitive skills.

- *Model D*: We intended to construct a SEM with all latent literacy variables as outcomes and all individual and environmental characteristics as predictors in order to investigate the effects of environmental and individual characteristics on literacy at 11.5 (without taking cognitive skills at 5.5 into account). Due to convergence issues in this preregistered model, we ran separate multivariate regressions for each literacy skill, in which each outcome was approximated by its factor score, extracted with the regression method, adapted to non-orthogonal factors and presenting maximal validity (DiStefano et al., 2019).

- *Model E*: We added to Model C all individual and environmental characteristics which had a significant effect on at least one literacy outcome in Model D as predictors, modelling the mediation relationships between individual/environmental characteristics and literacy outcomes via early cognitive and socio-emotional skills. Thus, the model estimated both direct paths (association of distal predictors, here environmental and individual variables, and outcomes, here literacy at 11.5) and indirect paths (association between distal predictors and outcomes mediated through mediators, here cognitive and socio-emotional abilities at 5.5). Correlations between all mediators were included in order to account for any unmodelled sources of covariation among mediators, and bias-corrected bootstrapped confidence intervals with 10,000 resamples were derived in order to account

for the likely non-normality of mediated effects (product of estimates), following recommendations by (Preacher & Hayes, 2008). This model was similar to the one in Chapter 3, Part 3.1.

Model estimation and inference. The models were estimated with the MLR estimator, given the non-normality of many of our variables, including some of our literacy outcomes. For all models, we checked modification indices to test for significant correlations between predictors, which were declared in the model when necessary. All models were fitted on the final sample as previously defined (N=297). The remaining missing data were treated using full information maximum likelihood (FIML) estimation. We examined measures of goodness-of-fit, including the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean squared error of approximation (RMSEA). CFI and TLI values greater than 0.95 and values of RMSEA less than 0.06 were used as cut-offs (following recommendations by Hu & Bentler, 1999). We reported fit indices corrected for non-normality: robust RSMEA as defined by Brosseau-Liard et al., (2012) and robust TLI and CFI as defined by Brosseau-Liard & Savalei (2014). We only interpreted standardized coefficients in order to estimate the relative influence of predictors on literacy outcomes. Significance of coefficients of interest in each model was assessed with two-sided t-tests – except for Model E where 95% bias-corrected bootstrapped confidence intervals were computed. In Models A-D, we accounted for multiple comparison by using the Benjamini-Hochberg False Discovery Rate (FDR) with the total number of tests equal to the number of hypotheses tested in each model, and a false discovery threshold $q=0.05$. For Model E, the main interest was the estimation of the standardized mediation effect; as such, we did not control for multiple testing. For Models A-D, standardized effects greater than 0.1 were considered meaningful. For Model E, standardized indirect effects greater than 0.05 were considered meaningful.

4.1.4 Results

Descriptive statistics

Descriptive statistics of cognitive, socio-emotional and literacy scores of included participants are presented in **Table 2** and **Table 3**.

Table 2*Descriptive statistics for cognitive and socio-emotional test scores at 5.5 years.*

Variable	N	Mean	SD	Min	Median	Max	Kurtosis	Skew
Non-word repetition (NEPSY, out of 46)	297	29.65	7.62	8.00	30.00	45.00	2.71	-0.44
Word-segment recognition (NEPSY, out of 14)	297	11.03	1.75	4.67	11.00	14.00	3.46	-0.64
Sentence repetition (NEPSY, out of 34)	297	16.69	4.00	8.00	16.00	28.00	2.86	0.25
Information (WPPSI-III, out of 34)	297	25.81	2.58	17.00	26.00	31.00	3.53	-0.51
Vocabulary (WPPSI-III, out of 43)	297	24.54	5.37	11.00	24.00	38.00	2.62	0.27
Word reasoning (WPPSI-III, out of 28)	297	17.23	4.26	0.00	18.00	27.00	3.86	-0.89
Peg moving task (time in seconds)	297	27.79	4.82	20.00	27.00	46.00	3.89	0.89
Design copying (NEPSY, out of 72)	292	48.01	6.68	31.20	47.51	67.09	2.76	0.29
Block design (WPPSI-III, out of 40)	296	29.21	3.63	20.00	30.00	40.00	3.44	0.07
Matrix reasoning (WPPSI-III, out of 29)	297	16.33	3.92	6.00	16.00	28.00	3.26	0.16
Picture concepts (WPPSI-III, out of 28)	297	15.00	3.77	5.00	15.00	24.00	2.54	-0.05
Coding (WPPSI-III, out of 65)	295	30.52	9.42	0.00	29.00	59.00	3.30	0.43
Emotional symptoms (SDQ, out of 10)	297	2.20	1.90	0.00	2.00	8.00	3.26	0.88
Conduct problems (SDQ, out of 10)	297	2.12	1.86	0.00	2.00	8.00	3.06	0.76
Hyperactivity/inattention (SDQ, out of 10)	297	2.74	2.33	0.00	2.00	10.00	3.33	0.84
Peer relationship problems (SDQ, out of 10)	297	1.09	1.17	0.00	1.00	6.00	4.75	1.23
Prosocial behaviour (SDQ, out of 10)	297	8.41	1.73	3.00	9.00	10.00	3.07	-0.99

Table 3*Descriptive statistics for literacy test scores at 11.5 years*

Variable	N	Mean	SD	Min	Median	Max	Kurtosis	Skew
Number of errors when reading text	297	1.09	2.37	0.00	0.00	20.00	30.91	4.69
Number of read words when reading text	297	146.41	27.7	52.00	148.00	214.00	3.07	-0.41
Orthographic choice (out of 13)	297	12.13	1.24	4.00	12.00	13.00	17.31	-3.07
Orthographic choice task time (in s)	297	2328.00	714.68	742.00	2160.70	5314.30	4.64	1.09
Orthography proportion correct (13 item)	297	0.00	0.09	-0.70	0.06	0.06	25.36	-3.87
Regular words reading time (in s)	297	20.46	6.16	6.00	19.00	54.00	6.48	1.33
Number of correctly read regular words (out of 20)	297	18.96	1.46	12.00	19.00	20.00	7.56	-2.02
Irregular words reading time (in s)	297	21.93	7.67	10.00	20.00	72.00	10.92	2.15
Number of correctly read irregular words (out of 20)	297	17.66	2.48	7.00	18.50	20.00	5.95	-1.66
Non-words reading time (in s)	297	27.86	8.38	13.00	26.00	68.00	7.04	1.48
Number of correctly read non-words (out of 20)	297	16.19	2.18	8.50	16.00	20.00	3.06	-0.59
Reading comprehension: first text (out of 12)	297	11.45	0.85	7.00	12.00	12.00	9.10	-2.12
Reading comprehension: second text (out of 12)	297	10.58	1.42	3.00	11.00	12.00	10.50	-2.11
Reading comprehension (out of 24)	297	22.03	1.77	11.00	22.00	24.00	9.30	-1.88
Accuracy score when spelling sentences (out of 20)	297	13.32	4.14	1.00	14.00	20.00	2.50	-0.38
Number of correctly spelt regular words (out of 20)	297	18.33	1.63	11.00	19.00	20.00	6.30	-1.55
Number of correctly spelt irregular words (out of 10)	297	7.98	1.76	2.00	8.00	10.00	3.36	-0.89
Number of correctly spelt non-words (out of 20)	297	15.99	2.73	0.00	16.00	20.00	9.41	-1.78

Specific early language predictors of literacy skills at 11.5

Figure 2 presents the structure of the CFA modelling the different literacy outcomes. Following the modification indices, correlations between each reading score and their corresponding response time were included, as well as one cross-loading between the latent reading accuracy variable and the number of words read in the timed text reading task (representing text reading speed). These correlations and cross-loading had the highest modification indices (>10) while being theoretically plausible. The high fit indices (RMSEA = 0.014; CFI = 0.997; TLI = 0.996) validated this measurement model. Overall, factor loadings were balanced, with small exceptions: spelling principally loaded on dictation tasks, and reading accuracy loaded on word-level variables more than on sentence-level ones. The four latent variables were highly correlated, reading speed being slightly less correlated with the other literacy skills.

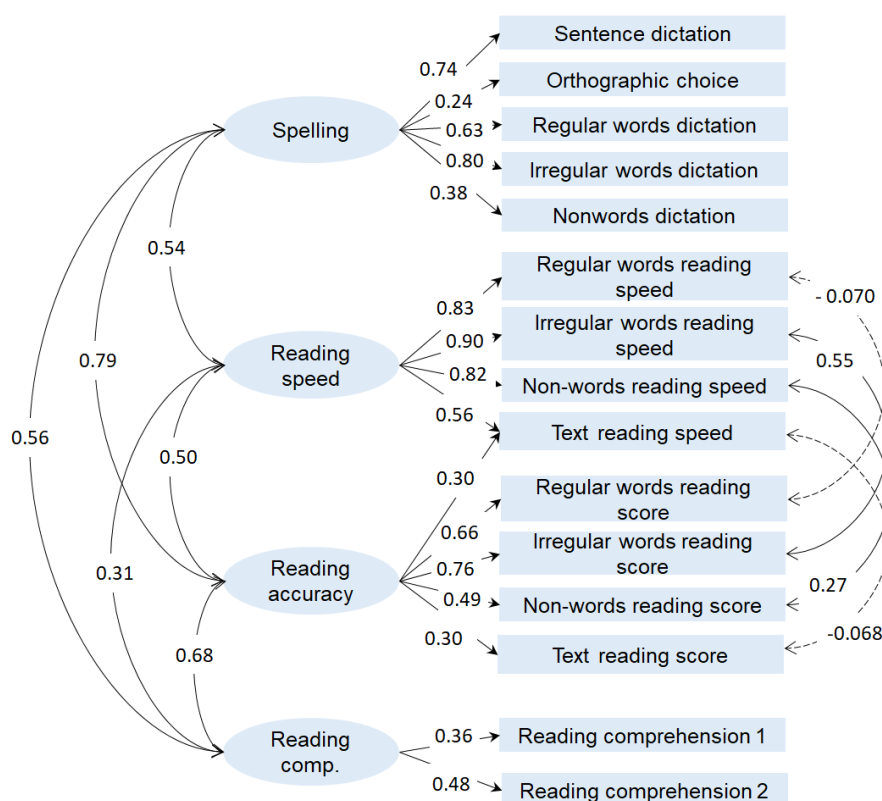


Figure 2: Final CFA of literacy skills at 11.5 years. Full line indicates significance; non-significant correlations are represented with a dotted line. All coefficients are standardized. $\chi^2(79, N=297) = 83.074$, p -value = 0.355. Robust Comparative Fit Index

(CFI) = 0.997. Robust Tucker-Lewis Index (TLI) = 0.996. Robust RMSEA = 0.014 [0.000; 0.038] (90% confidence interval).

Results from the longitudinal SEM with pre-school language abilities as predictors of the four literacy outcomes at 11.5 years old are presented in **Figure 3** and **Table 4**. Spelling was significantly predicted by sentence repetition and information, whereas reading accuracy was significantly predicted by non-word and sentence repetition (standardized β above 0.2 for the four relationships). Other standardized regression coefficients were non-significant. With a good fit (RMSEA = 0.031, CFI=0.972, TLI=0.962), this first model better explained reading comprehension ($R^2 = 0.29$), than reading accuracy ($R^2 = 0.20$) and spelling ($R^2 = 0.17$), but poorly explained reading speed ($R^2 = 0.05$).

Table 4

Language predictors of literacy skills at 11.5 years old.

	Spelling	Reading speed	Reading accuracy	Reading comprehension
	β (SD)	β (SD)	β (SD)	β (SD)
Predictors at 5.5 years old				
Non-word repetition	0.09 (0.08)	0.01 (0.07)	0.23 (0.08)*	0.10 (0.15)
Word-segment recognition	-0.05 (0.07)	0.06 (0.07)	-0.02 (0.07)	0.03 (0.14)
Sentence repetition	0.24 (0.09)*	0.15 (0.08)	0.20 (0.08)*	0.22 (0.21)
Information	0.24 (0.08)*	0.09 (0.09)	0.13 (0.07)	0.26 (0.13)
Vocabulary	-0.16 (0.09)	-0.01 (0.08)	-0.15 (0.08)	-0.24 (0.14)
Word reasoning	0.04 (0.09)	-0.03 (0.08)	0.10 (0.09)	0.24 (0.18)
Fit indices				
Chi-square	$\chi^2(145, N=297) = 185.12$ (p-value < 0.01)			
CFI	0.972			
TLI	0.962			
RMSEA [95% CI]	0.031 [0.015;0.044]			
R-square	0.169	0.045	0.196	0.291

Note: * indicates significance after FDR correction.

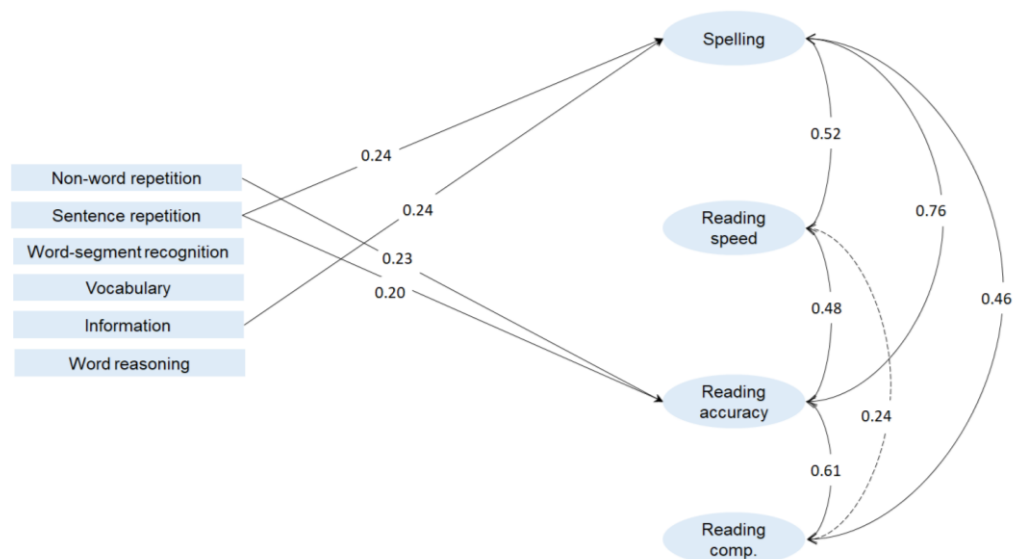


Figure 3: Early specific language predictors of literacy skills at 11.5. Only significant regressions coefficients are represented (after FDR correction for multiple comparison). As for correlations, a full line indicates significance; non-significant correlations are represented with a dotted line. All coefficients are standardized. For greater readability, factor loadings are not represented.

As a supplementary analysis, we ran univariate analyses corresponding to the Model A, regressing literacy outcomes on each language test at 5.5. Results of these analyses are available in **Table A2**. All preschool tests were positively and significantly associated with all literacy outcomes at 11.5, with the exception of word-segment recognition (which predicted none), and vocabulary (which only predicted reading accuracy).

Early cognitive and socio-emotional predictors of literacy skills at 11.5

Figure 4 presents the CFA of language, visuospatial and fine motor skills at 5.5 years. **Table 5** and **Figure 5** sum up the results of the longitudinal SEM including these skills as predictors of literacy outcomes. Language skills at 5.5 significantly predicted reading accuracy, reading comprehension and spelling at 11.5, with standardized β ranging from 0.27 for spelling to 0.41 for reading comprehension. Kindergarten visuospatial and fine motor skills did not significantly predict later literacy. The resulting model provided a good fit, the R-squares showing the same trend as for the previous model, with reading comprehension being best explained and reading speed being poorly explained.

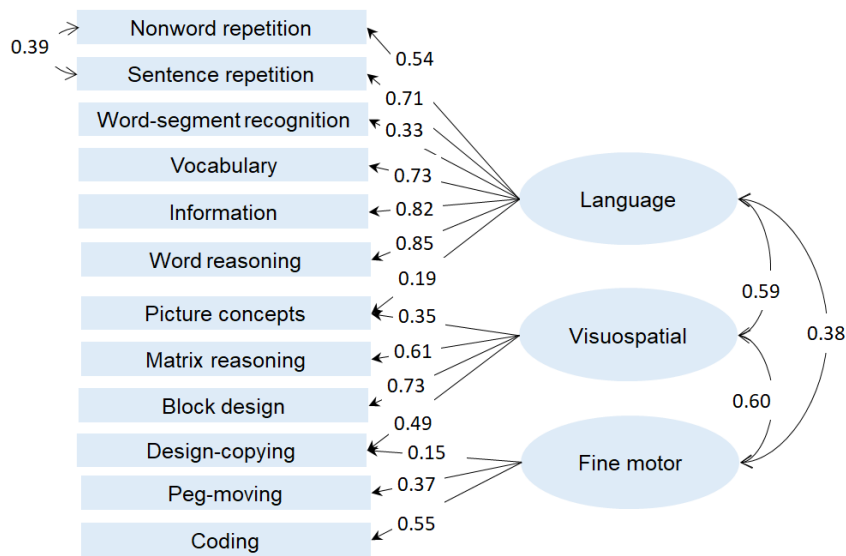


Figure 4: CFA of cognitive skills at 5.5 years. A full line indicates significance; non-significant correlations are represented with a dotted line. All coefficients are standardized. χ^2 (48, N=985) = 135.89 (p-value < 0.001). Robust Comparative Fit Index (CFI) = 0.974. Robust Tucker-Lewis Index (TLI) = 0.965. Robust RMSEA = 0.043 [0.035, 0.052] (90% confidence interval). On this figure, we present the coefficients as fitted on the sample at 5.5 years. For all model including this CFA, the whole model was fitted on the working sample (similar coefficients).

As a complementary analysis, we investigated the univariate associations between visuospatial and fine motor skills and literacy outcomes. Hence, we ran additional SEM models in literacy outcomes were regressed on each cognitive predictor. The results of these univariate analyses are presented in **Table A3**. We found that language, visuospatial and fine motor skills significantly predicted all literacy outcomes. In particular, standardized β ranged from 0.20 (reading speed) to 0.44 (reading comprehension) with visuospatial skills as predictor, and from 0.45 (reading speed) to 0.71 (reading comprehension) for fine motor skills.

Table 5*Cognitive predictors of literacy skills at 11.5 years old – with latent variables.*

	Spelling at 11.5	Reading speed at 11.5	Reading accuracy at 11.5	Reading comprehension at 11.5
	β (SD)	β (SD)	β (SD)	β (SD)
Predictors at 5.5				
Language skills at 5	0.27 (0.11)*	0.13 (0.1)	0.31 (0.12)*	0.41 (0.16)*
Visuo-spatial skills at 5	0.05 (0.27)	0.04 (0.24)	0.04 (0.24)	0.12 (0.24)
Fine motor skills at 5	0.30 (0.35)	0.29 (0.3)	0.22 (0.25)	0.07 (0.25)
Fit indices				
Chi-square	$\chi^2(295, N=297) = 361.96$ (p-value < 0.01)			
CFI	0.971			
TLI	0.965			
RMSEA [95% CI]	0.028 [0.016;0.037]			
R-square	0.210	0.125	0.187	0.255

Note: * indicates significance after FDR correction.

Figure 6 and **Table 6** present results from the longitudinal model including cognitive and socio-emotional variables as predictors of literacy outcomes. Language skills still significantly predicted reading accuracy, reading comprehension and spelling; however, the latter did not remain significant after correction for multiple comparisons. Hyperactivity at 5.5 negatively predicted reading accuracy at 11.5 (standardized $\beta = -0.21$) as well as spelling, although this latter effect did not reach significance after FDR correction. Likewise, the association between early emotional problems and reading accuracy and comprehension was not significant anymore after FDR correction. Adding socio-emotional variables resulted in a substantial increase in the part of variance explained of reading accuracy and reading comprehension, but not in that of reading speed and spelling.

As a supplementary analysis, we ran the same model splitting the hyperactivity/inattention score into hyperactivity and inattention. In this model, neither hyperactivity nor inattention predicted reading accuracy. Results and details are available in **Table A4**.

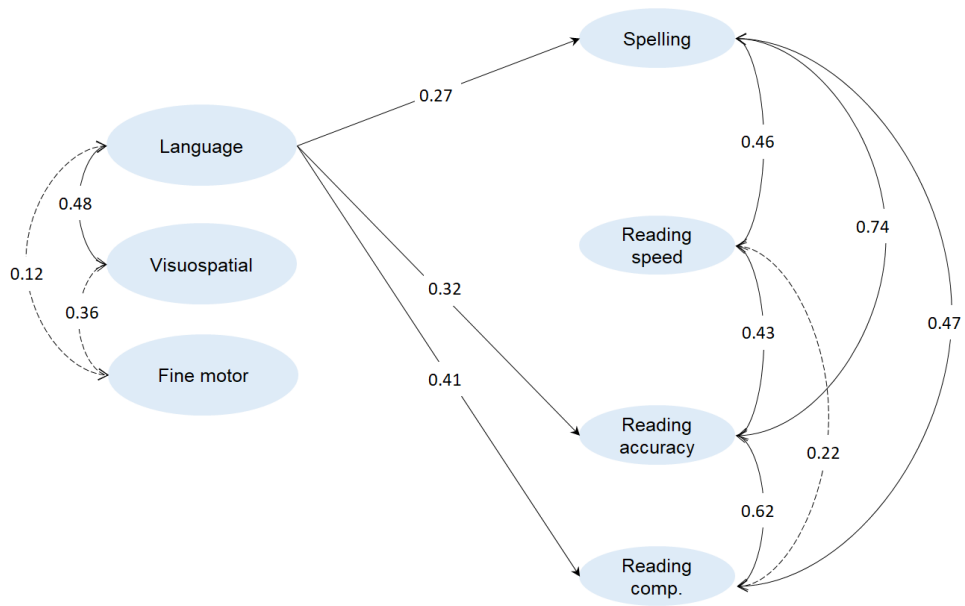


Figure 5: Early latent cognitive predictors of literacy skills at 11.5. Only significant regressions coefficients are represented (after FDR correction for multiple comparison). As for correlations, full line indicates significance; non-significant correlations are represented with a dotted line. All coefficients are standardized. For greater readability, factor loadings are not represented.

Table 6

Cognitive predictors of literacy skills at 11.5 years old – with latent cognitive abilities and socio-emotional variables.

Predictors	Spelling	Reading	Reading	Reading
	β (SD)	speed β (SD)	accuracy β (SD)	comprehension β (SD)
<i>Latent cognitive abilities</i>				
Language skills at 5	0.26 (0.10)	0.12 (0.09)	0.32 (0.11)*	0.47 (0.15)*
Visuo-spatial skills at 5	0.02 (0.19)	0.04 (0.17)	0.01 (0.20)	0.03 (0.22)
Fine motor skills at 5	0.29 (0.24)	0.27 (0.22)	0.23 (0.20)	0.13 (0.22)
<i>Behavioural/emotional abilities</i>				
Emotional symptoms	-0.10 (0.06)	-0.01 (0.06)	-0.13 (0.06)	-0.30 (0.14)
Conduct problems	-0.06 (0.09)	0.06 (0.08)	0.14 (0.09)	-0.01 (0.14)
Hyperactivity/inattention	-0.15 (0.07)	-0.11 (0.06)	-0.21 (0.07)*	-0.08 (0.13)
Peer relationship problems	0.05 (0.08)	-0.04 (0.08)	-0.07 (0.08)	-0.06 (0.12)
Prosocial behaviour	0.03 (0.07)	0.05 (0.07)	-0.02 (0.07)	-0.01 (0.13)
Fit indices				
Chi-square	$\chi^2(410, N=297) = 513.46$ (p-value < 0.01)			
CFI	0.956			
TLI	0.948			
RMSEA [95%CI]	0.029 [0.020;0.037]			
R-square	0.227	0.126	0.242	0.382

Note: * indicates significance after FDR correction.

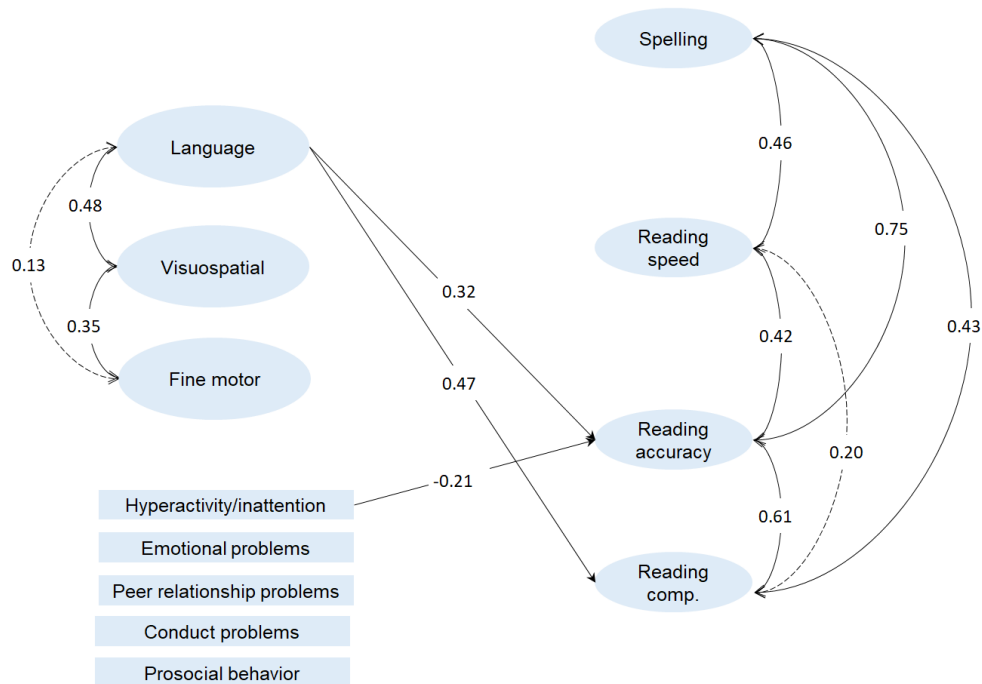


Figure 6: Early cognitive and socio-emotional predictors of literacy skills at 11.5. Only significant regressions coefficients are represented (after FDR correction for multiple comparison). As for correlations, a full line indicates significance; non-significant correlations are represented with a dotted line. All coefficients are standardized. For greater readability, factor loadings are not represented.

Pre- and post-natal predictors of literacy skills at 11.5

Results of the multiple regressions of environmental/individual factors on literacy outcomes are presented in **Table 7**. After correction for multiple comparisons, parental education strongly predicted reading accuracy and reading comprehension (standardized $\beta = 0.27$ and 0.27), and cognitive stimulation predicted reading accuracy (standardized $\beta = 0.17$). Before correction, breastfeeding negatively predicted reading speed (standardized $\beta = -0.15$); the number of siblings negatively predicted reading accuracy, spelling and reading comprehension (standardized $\beta = -0.13$, -0.16 and -0.14); and mother age positively predicted spelling and reading comprehension (standardized $\beta = 0.15$ and 0.12). R-squares followed the same trend as for previous models, but were substantially lower; the model only accounted for 10% of the variance for the best explained outcome, reading comprehension.

Table 7*Environmental and individual predictors of literacy skills at 11.5 years old.*

Predictors	Spelling		Reading speed	
	β (SD)	p-val.	β (SD)	p-val.
Sex (ref. = female)	-0.03 (0.06)	0.602	0.07 (0.06)	0.243
Gestational age	0.02 (0.06)	0.761	0.10 (0.06)	0.099
Birth weight	0.03 (0.07)	0.674	0.00 (0.07)	0.974
Mother age	0.15 (0.07)	0.030	0.08 (0.07)	0.275
Number of older siblings	-0.16 (0.07)	0.020	-0.08 (0.07)	0.254
Parental education	0.22 (0.08)	0.004	0.20 (0.08)	0.01
Household monthly income	0.00 (0.08)	0.967	0.00 (0.08)	0.953
Breastfeeding duration (1=more than 3 days)	-0.14 (0.06)	0.024	-0.15 (0.06)	0.016
Cognitive stimulation	0.13 (0.06)	0.042	0.11 (0.06)	0.104
Maternal depression episode before 5.5 years old	-0.02 (0.06)	0.722	0.03 (0.06)	0.570
Alcohol consumption during pregnancy	-0.08 (0.06)	0.163	-0.07 (0.06)	0.272
Tobacco consumption during pregnancy	0.02 (0.06)	0.797	-0.01 (0.06)	0.908
R-square	0.0695		0.0393	
Predictors	Reading accuracy		Reading comprehension	
	β (SD)	p-val.	β (SD)	p-val.
Sex (ref. = female)	-0.01 (0.06)	0.910	-0.03 (0.05)	0.484
Gestational age	0.05 (0.06)	0.394	0.05 (0.05)	0.278
Birth weight	0.05 (0.07)	0.437	0.02 (0.06)	0.718
Mother age	0.11 (0.07)	0.117	0.12 (0.06)	0.048
Number of older siblings	-0.13 (0.07)	0.041	-0.14 (0.06)	0.013
Parental education	0.27 (0.08)*	<0.001	0.27 (0.06)*	<0.001
Household monthly income	-0.03 (0.08)	0.732	-0.03 (0.07)	0.607
Breastfeeding duration (1=more than 3 days)	-0.11 (0.06)	0.074	-0.05 (0.05)	0.329
Cognitive stimulation	0.17 (0.06)*	0.006	0.14 (0.05)	0.009
Maternal depression episode before 5.5 years old	0.01 (0.06)	0.875	-0.03 (0.05)	0.506
Alcohol consumption during pregnancy	-0.09 (0.06)	0.116	-0.06 (0.05)	0.255
Tobacco consumption during pregnancy	0.03 (0.06)	0.620	0.03 (0.05)	0.611
R-square	0.0762		0.101	

Note: * indicates significance after FDR correction. Adjusted for centre.

Mediation of pre-and post-natal factors effects through early cognitive and socio-emotional skills

Parental education and cognitive stimulation, which were the only factors to have a significant effect after correction in Model D, were included as distal predictors in a mediation model with cognitive and socio-emotional skills at 5.5 as mediators, and latent literacy skills as outcomes. For each environmental variable and literacy skill, we were interested in the following effects, presented in **Table 8**:

- specific indirect effects, which are associations between the predictor and the literacy outcome mediated through a given mediator;
- total indirect effects, the sum of indirect effects on all possible mediators;
- direct effects, which are associations between the predictor and the outcome which is not explained by indirect effects;
- total effects, the sum of total indirect and direct effects.

The significant effects are summed up in **Figure 7**. Parental education had a significant total effect on all literacy outcomes at 11.5, ranging from a standardized effect of 0.41 (95% bias-corrected bootstrapped C.I. = [0.01, 28.8]) for reading comprehension, to 0.2 [0.04; 1.06] for reading speed. For reading accuracy and spelling, the effect of parental education was significantly mediated through language skills at 5.5 (standardized β of 0.10 [0.01; 1.74] and 0.05 [-0.02; 5.93] respectively). Cognitive stimulation only had a significant total effect on reading accuracy, although non-significant, indirect effects of cognitive stimulation on reading speed, spelling and reading accuracy through fine motor skills were large (standardized $\beta \geq 0.10$). This final model explained 41.5% of variance in reading comprehension, 30.6% of variance in reading accuracy, 29.0% of variance in spelling, and 22.7% of variance in reading speed at 11.5. Compared to Model C, adding environmental and individual predictors decreased the global fit of the model, which remains acceptable (RMSEA= 0.030, CFI=0.957, TLI=0.944).

Table 8

Mediation model with the significant environmental/individual predictors identified in model D.

	Reading accuracy		Reading speed	
	Parental education	Cognitive stimulation	Parental education	Cognitive stimulation
	β [95% B.C.C.I]	β [95% B.C.C.I]	β [95% B.C.C.I]	β [95% B.C.C.I]
Parameters				
Total	0.29 [0.12;0.51]	0.22 [0.03;0.49]	0.20 [0.04;1.06]	0.07 [-0.08;0.32]
Total indirect	0.10 [-12.8;0.24]	0.13 [-0.02;23.61]	0.00 [-24.28;0.13]	0.15 [-0.01;36.52]
Direct effect	0.19 [-0.02;10.21]	0.08 [-11.88;0.32]	0.20 [0.00;20.39]	-0.08 [-32.8;0.12]
<i>Indirect effect through</i>				
Language skills	0.10 [0.01;1.74]	0.05 [0.00;1.23]	0.05 [-0.02;5.93]	0.02 [-0.01;3.72]
Visuo-spatial skills	-0.03 [-6.60;0.10]	0.00 [-0.03;1.07]	-0.03 [-10.53;0.06]	0.01 [-0.02;2.92]
Fine motor skills	-0.02 [-15.11;0.09]	0.10 [-0.01;24.16]	-0.03 [-19.49;0.15]	0.13 [0.00;34.92]
Emotional symptoms	0.01 [-0.03;0.23]	-0.01 [-0.83;0.01]	0.00 [-0.01;2.09]	-0.01 [-3.00;0.00]
Conduct problems	0.00 [-0.03;0.20]	-0.01 [-0.42;0.01]	0.00 [-0.02;0.62]	0.00 [-1.70;0.01]
Hyperactivity/inattention	0.03 [-0.12;0.13]	0.00 [-0.02;0.12]	0.01 [-1.50;0.06]	0.00 [-0.02;0.17]
Peer relationship problems	0.00 [-0.02;0.27]	0.00 [-0.02;0.66]	0.00 [-0.21;0.05]	0.00 [-0.15;0.07]
Prosocial behaviour	0.00 [-0.27;0.02]	0.00 [-1.08;0.01]	0.00 [-0.02;0.38]	0.00 [-0.04;0.35]
Fit indices				
R-square	0.306		0.227	

Table 8 (continued)

Parameters	Spelling		Reading comprehension	
	Parental education	Cognitive stimulation	Parental education	Cognitive stimulation
	β [95% B.C.C.I]	β [95% B.C.C.I]	β [95% B.C.C.I]	β [95% B.C.C.I]
Total	0.24 [0.07;1.28]	0.07 [-0.15;0.71]	0.41 [0.01;28.80]	0.18 [-0.85;17.93]
Total indirect	0.06 [-19.23;0.22]	0.15 [-0.02;36.02]	0.14 [-14.17;8.98]	0.11 [-2.88;24.36]
Direct effect	0.19 [-0.04;15.12]	-0.09 [-25.88;0.16]	0.27 [-1.86;29.41]	0.06 [-16.81;17.65]
<i>Indirect effect through</i>				
Language skills at 5	0.09 [0.01;5.29]	0.05 [-0.01;2.84]	0.15 [-1.35;9.48]	0.07 [-0.44;5.92]
Visuo-spatial skills at 5	-0.03 [-14.36;0.09]	0.01 [-0.03;5.72]	-0.01 [-10.59;5.32]	0.00 [-1.52;2.34]
Fine motor skills at 5	-0.02 [-20.14;0.13]	0.12 [-0.02;33.74]	-0.01 [-30.77;0.13]	0.06 [-0.23;35.79]
Emotional symptoms	0.00 [-0.03;1.05]	-0.01 [-1.94;0.01]	0.01 [-0.79;3.10]	-0.02 [-4.26;0.31]
Conduct problems	0.00 [-0.56;0.04]	0.00 [-0.04;1.40]	0.00 [-1.02;0.56]	0.00 [-0.51;1.67]
Hyperactivity/inattention	0.02 [-1.69;0.09]	0.00 [-0.03;0.45]	0.00 [-2.17;4.11]	0.00 [-0.34;1.19]
Peer relationship problems	-0.01 [-0.63;0.03]	-0.01 [-0.86;0.03]	0.00 [-0.25;2.68]	0.00 [-0.21;2.93]
Prosocial behaviour	0.00 [-0.20;0.11]	0.00 [-1.00;0.04]	0.00 [-1.49;0.44]	0.00 [-2.89;0.16]
Fit indices				
Chi-square	$\chi^2(435, N=290) = 549.56$ (p.val < 0.01)			
CFI	0.957			
TLI	0.944			
RMSEA [95%CI]	0.030 [0.022;0.038]			
R-square	0.290		0.415	

Note: Coefficients in bold are significant.

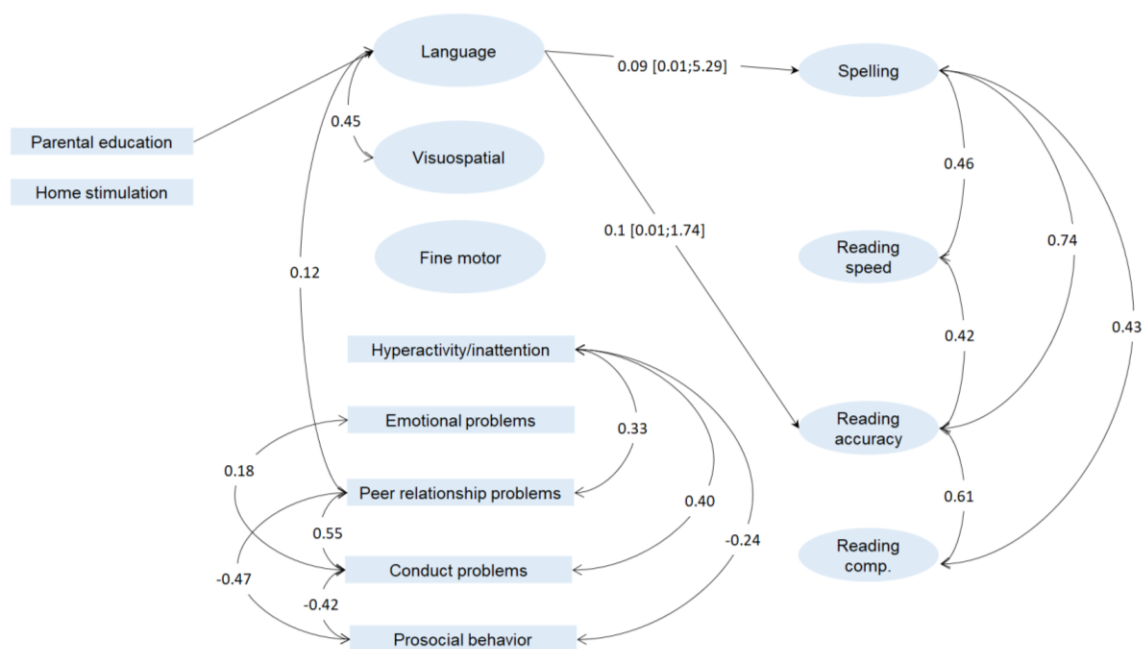


Figure 7: Mediation model. Only significant regressions coefficients are represented. In brackets are the corresponding bias-corrected bootstrapped confidence intervals (BCCI) with 10,000 resamples. As for correlations, a full line indicates significance. All coefficients are standardized. For greater readability, factor loadings are not represented.

4.1.5 Discussion

In this study, we investigated the relative influence of a broad range of preschool cognitive, socio-emotional, individual and environmental factors on reading speed, reading accuracy, reading comprehension, and spelling at the end of primary school. Furthermore, we integrated these factors into a mediation model where individual and environmental characteristics were modelled as distal factors, assessing their direct and indirect effects on literacy skills through preschool cognitive and socio-emotional skills, included as more proximate factors.

Which pre-school language abilities are the best unique predictors of literacy skills at 11.5 years old?

Our results show that non-word repetition and sentence repetition at 5.5 were the best predictors of reading accuracy at 11.5, emphasizing the widely-accepted role of

phonological processing and short-term memory as fundamental building blocks of reading accuracy (Hjetland et al., 2020; Melby-Lervåg et al., 2012). More surprisingly, word segment recognition, which also taps phonological processing, and in particular, phonemic awareness, did not predict reading accuracy – nor any other literacy outcome, even in univariate analyses –, although phonemic awareness has been shown to be the best phonological predictor of word recognition (Melby-Lervåg et al., 2012). This may be due to this test's low reliability in our sample (Cronbach's $\alpha = 0.43$) compared with that of the other language tests. Early vocabulary was not significantly associated with later reading accuracy, in contrast with previous results in a large Chinese sample (Su et al., 2017), but consistent with previous results from French and British samples (Muter et al., 2004; Piquard-Kipffer & Sprenger-Charolles, 2013b). On the other hand, reading speed was not significantly predicted by any early language test in multivariate analyses. It is possible that we found no effect from the phonological tasks because they did not tap into the speed of access to phonological representations, as naming speed has been shown to be the best phonological predictor of reading speed in French (Moll et al., 2014). Similarly, and unexpectedly, no predictor was significantly associated with reading comprehension, in contrast with meta-analytic results pointing to an influence of vocabulary, grammar, as well as phonological awareness (Hjetland et al., 2020). Again, the low reliability of our reading comprehension tests may be the culprit here. Turning to spelling performance, we found significant associations with verbal short-term memory, as assessed by sentence repetition. This result is consistent with Moll et al. (2014), who identified phonological short-term memory as the best predictor for spelling in French, above phonological awareness. The information test from the WPPSI-III, tapping language comprehension and concept knowledge, also significantly predicted spelling performance, which had not been observed previously in the literature – a result which would thus need to be replicated.

How do preschool general language skills, visual-spatial skills and fine-motor skills differentially predict acquired literacy skills?

As expected, early language predicted reading accuracy, reading comprehension and spelling more than other cognitive skills; only reading speed was unexpectedly not predicted by language at 5.5. Visuospatial and fine motor skills did not predict any literacy outcome beyond language skills in our study. These results are in line with those from Piquard-Kipffer & Sprenger-Charolles (2013), who showed no effect of kindergarten visual short-term memory on later reading fluency, but contrast with studies which used specific

visual attention measures (Bosse et al., 2015; Bosse & Valdois, 2009; Plaza & Cohen, 2007; Valdois et al., 2019). It is likely that the general visuospatial skills latent variable that we used correlates less with reading skills than measures of visual attention do. The non-significant effect of fine-motor skills on literacy abilities is consistent with results from Pitchford et al. (2016), who found that fine-motor abilities do not predict reading when short-term memory is controlled. However, it is possible that this absence of effect is partly explained by the fact that we measured literacy outcomes at 11.5, as previous studies have shown effects on very early reading (Cameron et al., 2012; Grissmer et al., 2010; Suggate et al., 2018, 2019).

Do certain early socio-emotional difficulties/abilities predict later literacy skills, beyond these early cognitive skills?

As predicted, early inattention-hyperactivity symptoms had a negative effect on reading outcomes; in particular, reading accuracy (and, marginally, on spelling). The absence of effect on reading speed and reading comprehension is more surprising, but to our knowledge, no study has compared the impact of early socio-emotional variables on a wide range of literacy outcomes. Few long-term longitudinal studies investigated the impact of socio-emotional symptoms on reading in non-clinical populations (Halonen et al., 2006), and several found no effect of externalizing and internalizing problems on reading achievement once attention was controlled for (Breslau et al., 2009b; Duncan et al., 2007; Rabiner & Coie, 2000). Our results are consistent with these studies, since the only SDQ variable that had significant effect on literacy outcomes is the one including a measure of inattention. When distinguishing inattention and hyperactivity, neither predicted reading accuracy; however, our inattention measure (2 items) was probably too limited to replicate the inattention effect found in the literature. Emotional problems also marginally predicted some literacy outcomes, which is in line with results from Carroll et al. (2005), Massetti et al. (2008) and Willcutt & Pennington (2000) (but see Breslau et al., 2009; Duncan et al., 2007; Rabiner & Coie, 2000). Therefore, behavioral and emotional problems displayed prior to reading instruction are negatively associated with later literacy outcomes and thus may be hampering reading acquisition. However, these influences are likely to be mutual, as reading development and socio-emotional symptoms seem to influence each other throughout the years of schooling (Halonen et al., 2006).

Which pre- and post-natal individual and environmental factors uniquely predict literacy skills at 11.5 years old?

Parental education was positively associated with reading accuracy and comprehension, which was consistent with our hypothesis and the literature (Sirin, 2005; Su et al., 2017). Unexpectedly, parental education did not exert any significant influence on reading speed and spelling, contrasting with results from Su et al. (2017). Cognitive stimulation also positively predicted reading accuracy at 11.5, in line with Belsky et al. (2007) and Sammons et al. (2015), who found long-term effects of the home literacy environment (12 years and beyond) even when controlling for late parental practices. These findings advocate for an early role of cognitive stimulation. By contrast, no effect of sex, number of siblings, breastfeeding and other pre- and post-natal risk factors were found. The absence of difference between girls and boys in literacy outcomes is particularly surprising, as the gender gap in reading comprehension and writing is firmly established in the literature (Lynn & Mikk, 2009; Reilly et al., 2018). This result contrasts with the very large sex difference ($\beta = 0.5$) that we found in favour of boys in arithmetic problem solving in the same sample, at the same age (Chapter 3, Part 3.1). Overall, one can note that environmental and individual predictors explained a small part of the variance in literacy outcomes compared to previous models including cognitive predictors, in line with Elbro & Scarborough (2004).

To what extent are the effects of parental education and cognitive stimulation mediated through early cognitive and behavioral skills?

Integrating these factors into a mediation model, we found that the effect of parental education on literacy outcomes (reading accuracy and spelling) was significantly mediated through early language skills. This result is consistent with Su et al. (2017) and Zhang et al. (2013), who found that the effect of SES on reading outcomes was fully mediated by early literacy in samples of Chinese children. Note however that in our study, neither total direct nor total indirect effects were significant, though both were large. Replication with a larger sample size would be needed to conclude on whether parental education exerts influences on literacy skills solely through early effects on the child's language skills, or also has later persistent effects throughout schooling, for example by providing an adequate environment supporting academic achievement.

Chapter 4 - Early predictors of literacy skills

The effect of early cognitive stimulation seemed mostly mediated through early cognitive skills, although no indirect nor direct effect appeared significant (the sole significant effect was the total effect on reading accuracy). According to Senechal & LeFevre (2002), home stimulation can be decomposed into formal (parent teaching) and informal stimulation (storybook exposure) – the first promotes the acquisition of early literacy skills whereas the latter has effect on oral language skills. Investigating different age ranges, Hood et al. (2008), Lehl et al. (2020), Niklas & Schneider (2013) and Senechal (2006) indeed found partial to total mediation effect of cognitive stimulation on different literacy outcomes through early language and literacy skills. Our results seem consistent with this literature.

Limitations

These results must be read with the following limitations in mind. First, our sample size ($N=297$), while relatively large compared to the existing literature (in their review, Hjetland et al. (2020) pointed out that 44 studies out of 64 had a sample size inferior to 150), is still limited, which may explain the scarcity of significant results in such complex models. Moreover, the original sample of EDEN is not representative of the French population: urban, well-educated and high-income households are over-represented compared with the national population, and illiterate mother were not included. Selective attrition amplified this trend, as showed in **Table A1**. Participants present at the 11.5 year-old wave significantly tended to come from wealthier and more educated families, had less conduct problems and hyperactivity symptoms, and had higher scores on many cognitive tests compared to participants present at 5.5 years.

Second, as the data at 5.5 years old was not specifically collected for the present study, our measures have several shortcomings. Indeed, we could not include a set of frequently used predictors of literacy skills: rapid naming (Elbro & Scarborough, 2004; Hjetland et al., 2020; National Early Literacy Panel, 2008b), letter knowledge (Elbro & Scarborough, 2004; Hjetland et al., 2020; National Early Literacy Panel, 2008b), grammar (Hjetland et al., 2020), rime awareness (Hjetland et al., 2020; Muter et al., 2004), among others. The lack of RAN and letter knowledge is especially limiting, as they are considered to be major early predictors of later decoding and spelling skills. Furthermore, some of our tests have limited quality: 7 of our tests had a Cronbach's α below 0.7. Word-segment recognition at 5.5 and reading comprehension tests at 11.5 had particularly low reliability

values (α below 0.6). Lastly, one can also note that, even if we labelled several predictors as environmental, we cannot, without appropriate controls for genetic factors, rule out that these effects could be partly or fully of genetic origin. Relatedly, it is worth reminding that such a correlational study is not enough to deduce causality, thus the need for interventional studies to go further simple associations – even though the causal role of some of our predictors (e.g., phonological awareness and vocabulary; Elbro & Scarborough, 2004) has already been demonstrated through such interventions.

Strengths

Despite these limits, we contribute to the existing literature with a study including a variety of early predictors, allowing us to test multiple hypotheses and isolate the effects of a wide set of intertwined cognitive, socio-emotional, environmental and individual factors. We studied their influence on four different literacy outcomes, including the widely-studied reading accuracy and comprehension, but also spelling and reading speed, on which studies are rare. As advocated for in the literature (Kendeou et al., 2009; Moll et al., 2014; Su et al., 2017), our longitudinal design spanned a particularly wide age range, beginning shortly before reading instruction (which is critical to understand the origins of individual differences in reading acquisition) and reaching the end of primary school (closer to representing adult proficiency than usual studies in primary school). Finally, we investigated the role of early predictors of reading in French, a language with different orthographic depth than English, which has never been done with this sample size, variety of predictors and outcomes, and long-term follow-up.

These contributions are backed by a robust analysis. As advocated by Hjetland et al. (2020) in their review, our literacy tests were standardized and commonly used; and our cognitive tests at 5.5 years mostly had a large number of items and good reliability. We analysed our data with structural equation modelling, a data analysis technique controlling for measurement error, its use being promoted by Hjetland et al. (2020) but still rare in the literature. We checked and reported non-normality of our variables when necessary, while ensuring the analysis would not be compromised by it. We drew our main conclusions on significance corrected for multiple comparisons, which is not done in most studies, and might contribute to explain why we find fewer significant associations than what is reported in the literature. Finally, we studied continuous socio-emotional variables, like Breslau et al. (2009) and Duncan et al. (2007) did, which avoided issues encountered by Masseti et

al. (2008) and Rabiner & Coie (2000), who drew conclusions on very small sample of ADHD students, as they studied socio-emotional variables dichotomously.

Conclusion

Our results shed a new light on the early predictors of literacy acquisition in French. Incorporating a large variety of predictors, this study confirms the fundamental role of early language abilities – especially phonological ones – on literacy outcomes, overshadowing that of other early cognitive skills. These results also corroborate the detrimental influence of early socio-emotional symptoms (in particular, hyperactivity/inattention and emotional disorders) on later reading skills, emphasizing the importance of screening and remediating for these symptoms as early as possible. Lastly, our findings confirm the widely-established impact of parental education and home cognitive stimulation on children reading abilities, which eclipsed the role of other environmental predictors. Combining these factors in a mediation model, we provided further elements on the mechanisms by which parental education and cognitive stimulation influenced reading skills, and found their effects to be at least partly mediated through early language abilities. Focusing on language skills as early as possible in schools could thus be a means to support later literacy abilities in children from disadvantaged environments, and thus reduce the SES gap in reading.

4.1.6 Appendix

Table A1*Descriptive statistics comparing included and excluded participants*

Variable	Present at 11.5 years			Present at 5.5 years			Difference
	N	Mean	SD	N	Mean	SD	P-value
<i>Environmental/individual characteristics</i>							
Mother age (years)	297	29.87	4.72	985	29.76	4.77	0.7256
Number of older siblings	297	0.7	0.87	985	0.79	0.93	0.1248
Parental education (years)	297	14.35	2.15	985	13.53	2.28	<0.0001
Household monthly income (k€)	297	2936.48	907.38	985	2683.74	966.2	<0.0001
Breastfeeding duration (1=more than 3 days)	297	0.78	0.41	985	0.72	0.45	0.0312
Cognitive stimulation (HOME score, out of 27)	290	17.42	2.12	950	17.27	2.3	0.3019
Maternal depression episode before 5.5 years old (1=yes)	297	0.33	0.47	985	0.33	0.47	1.0000
Alcohol consumption during pregnancy (nb drinks/week)	297	0.54	1.47	984	0.58	1.49	0.6822
Tobacco consumption during pregnancy (1=yes)	297	0.13	0.34	984	0.22	0.41	0.0002
Sex (1=male)	297	0.48	0.5	985	0.53	0.5	0.1315
Gestational age (weeks)	297	39.31	1.6	985	39.27	1.73	0.7112
Birth weight (kg)	297	3.33	0.46	985	3.29	0.51	0.2011
Centre							
<i>Cognitive abilities at 5</i>							
Non-word repetition (NEPSY, out of 46)	297	29.65	7.62	985	28.29	7.98	0.0079
Word-segment recognition (NEPSY, out of 14)	297	11.03	1.75	985	10.94	1.9	0.4468
Sentence repetition (NEPSY, out of 34)	297	16.69	4	985	15.42	4.09	<0.0001
Information (WPPSI-III, out of 34)	297	25.81	2.58	985	25.08	2.91	<0.0001
Vocabulary (WPPSI-III, out of 43)	297	24.54	5.37	985	23.66	5.6	0.0146
Word reasoning (WPPSI-III, out of 28)	297	17.23	4.26	985	16.17	4.69	0.0003
Peg moving task (time in seconds)	297	27.79	4.82	982	28.15	5.18	0.2683
Design copying (NEPSY, out of 72)	292	48.01	6.68	971	46.49	6.91	0.0008
Block design (WPPSI-III, out of 40)	296	29.21	3.63	983	28.17	3.84	<0.0001
Matrix reasoning (WPPSI-III, out of 29)	297	16.33	3.92	985	15.39	3.91	0.0003
Picture concepts (WPPSI-III, out of 28)	297	15	3.77	984	14.33	3.97	0.0083

Chapter 4 - Early predictors of literacy skills

Coding (WPPSI-III, out of 65)	295	30.52	9.42	979	29.86	9.85	0.2971
Emotional symptoms (SDQ, out of 10)	297	2.2	1.9	985	2.12	1.87	0.5235
Conduct problems (SDQ, out of 10)	297	2.12	1.86	985	2.36	2.05	0.0577
Hyperactivity/inattention (SDQ, out of 10)	297	2.74	2.33	985	3.14	2.39	0.0102
Peer relationship problems (SDQ, out of 10)	297	1.09	1.17	985	1.23	1.28	0.0777
Prosocial behaviour (SDQ, out of 10)	297	8.41	1.73	985	8.36	1.68	0.6605

Table A2

Univariate analysis of language predictors of literacy skills at 11.5 years old.

	Spelling	Reading speed	Reading accuracy	Reading comprehension
	β (SD)	β (SD)	β (SD)	β (SD)
Predictors at 5.5 years old				
Non-word repetition	0.25 (0.06)*	0.12 (0.06)*	0.36 (0.06)*	0.31 (0.11)*
Word-segment recognition	0.04 (0.07)	0.1 (0.07)	0.08 (0.07)	0.13 (0.15)
Sentence repetition	0.35 (0.06)*	0.19 (0.06)*	0.37 (0.06)*	0.42 (0.16)*
Information	0.32 (0.06)*	0.15 (0.07)*	0.28 (0.07)*	0.41 (0.1)*
Vocabulary	0.13 (0.07)	0.11 (0.06)	0.16 (0.07)*	0.18 (0.09)
Word reasoning	0.25 (0.07)*	0.12 (0.06)*	0.29 (0.07)*	0.42 (0.13)*

Note: * indicates significance (p-value \geq 0.05).

Table A3

Univariate analysis of cognitive predictors of literacy skills at 11.5 years old.

	Spelling	Reading speed	Reading accuracy	Reading comprehension
	β (SD)	β (SD)	β (SD)	β (SD)
Predictors at 5.5 years old				
Language skills	0.34 (0.07)*	0.19 (0.07)*	0.37 (0.07)*	0.49 (0.13)*
Visuo-spatial skills	0.38 (0.08)*	0.26 (0.07)*	0.36 (0.11)*	0.44 (0.17)*
Fine motor skills	0.66 (0.17)*	0.45 (0.13)*	0.62 (0.18)*	0.71 (0.31)*

Note: * indicates significance (p-value \geq 0.05).

Table A4

Cognitive predictors of literacy skills at 11.5 years old – with latent cognitive abilities and socio-emotional variables differentiating hyperactivity and inattention.

	Spelling	Reading speed	Reading accuracy	Reading comprehension
	β (SD)	β (SD)	β (SD)	β (SD)
Predictors				
<i>Latent cognitive abilities</i>				
Language skills at 5	0.26 (0.1)	0.12 (0.09)	0.33 (0.11)*	0.46 (0.15)*
Visuo-spatial skills at 5	0.03 (0.16)	0.04 (0.15)	-0.01 (0.19)	0.04 (0.21)
Fine motor skills at 5	0.28 (0.20)	0.26 (0.18)	0.24 (0.18)	0.13 (0.20)
<i>Behavioural/emotional abilities</i>				
Emotional symptoms	-0.09 (0.06)	-0.02 (0.06)	-0.15 (0.06)	-0.31 (0.14)
Conduct problems	-0.07 (0.09)	0.04 (0.08)	0.16 (0.09)	0.02 (0.14)
Hyperactivity	-0.06 (0.09)	0.00 (0.08)	-0.16 (0.08)	-0.15 (0.15)
Inattention	-0.12 (0.08)	-0.13 (0.08)	-0.09 (0.09)	0.06 (0.12)
Peer relationship problems	0.06 (0.08)	-0.04 (0.08)	-0.08 (0.08)	-0.05 (0.12)
Prosocial behaviour	0.02 (0.07)	0.04 (0.07)	-0.01 (0.07)	0.00 (0.13)
Fit indices				
Chi-square	$\chi^2(433, N=297) = 540.052$ (p-value < 0.01)			
CFI	0.955			
TLI	0.946			
RMSEA [95% CI]	0.029 [0.020;0.037]			
R-square	0.221	0.125	0.252	0.391

Note: * indicates significance after FDR correction. The hyperactivity score is coded as the sum of three items indicating hyperactivity or impulsivity (max score = 6). The inattention score is coded as the sum of two items indicating inattention (max score= 6).

Chapter 5 - Sex differences in academic skills

5.1 Sex differences are modulated by evaluation type

This section is based on the following article:

Guez, A., Peyre, H., & Ramus, F. (*under review*). Sex differences in academic achievement are modulated by evaluation type.

5.1.1 Abstract

Studies on sex differences in academic skills have often reported diverging results depending on the type of evaluation used, with girls typically obtaining better school grades and results at national examinations, and boys scoring higher at standardized tests. In this paper, we provide a framework for better understanding and interpreting these differences, integrating previously established factors that affect variations in the gender gap across evaluation types: writing skills, stress, self-discipline and grading bias. We apply this framework to a dataset containing the results of 23,451 French students in three evaluations characterized by different combinations of these factors: teacher evaluations, national examinations, and standardized tests. Our analyses show that the direction and size of sex differences depend both on the subject (French or mathematics) and on how skills are measured. Furthermore, our results offer new insights regarding the extent to which each factor may influence the observed gap.

Keywords: Sex differences; Evaluation characteristics; Standardized tests; National examinations; Teacher grades

5.1.2 Introduction

The question of sex differences in academic skills has garnered much attention and concern from researchers and policy-makers in the last decades. However, studies on the topic often report diverging results depending on how these skills are measured, which muddles interpretation.

On the one hand, several meta-analyses analyzing differences in achievement test scores have shown that girls obtain better results at language tests (Hedges & Nowell, 1995; Hyde & Linn, 1988), while boys perform better in mathematics (Else-Quest et al., 2010; Hyde et al., 1990; Reilly et al., 2015), albeit not consistently (Lindberg et al., 2010). The much publicized PISA (Programme for International Student Assessment) studies have confirmed these findings: at age 15, in most participating countries, girls outperformed boys in reading assessments, while the gap was reversed in mathematics (OECD, 2015). On the other hand, studies focusing on school marks and examinations consistently reported an advantage of girls in all subjects. Thus, the meta-analysis by Voyer and Voyer (2014) revealed that girls outperformed boys across all course materials, the largest difference being in language and the smallest in mathematics. Similarly, in a large sample of UK students, Deary, Strand, Smith, and Fernandes (2007) found that girls performed better than boys in all subjects of the GCSE (with the exception of Physics where there was no difference). Several factors can help explain this apparent contradiction between results stemming from achievement tests and school evaluations.

First, some evaluations are more likely to assess certain aspects of students' personality and behavior, and in particular, *self-discipline*. In this context, self-discipline is defined as the ability to make a conscious effort to resist impulses in order to reach a higher goal (Duckworth & Seligman, 2006). As such, it is likely to affect teacher evaluation grades, since they reflect students' behavior in class and homework assignment completion. Previous research has shown that girls tend to display higher levels of self-discipline than boys (e.g., meta-analysis of gratification delay tasks by Silverman, 2003), and this difference partly explains why girls obtain better report card grades than predicted by their achievement test scores (Duckworth et al., 2015; Duckworth & Seligman, 2006; Kling et al., 2013). Similarly, Steinmayr and Spinath (2008) found that girls' lesser tendency to avoid work (as well as greater agreeableness) partly explained girls' advantage in German school marks (but not in mathematics).

Second, some evaluations tend to assess *writing skills* more than others, depending on the type of response used (typically, free response versus multiple choice questions). Since girls display higher writing abilities than boys (Feingold, 1988; Hedges & Nowell, 1995; Reilly et al., 2018; Scheiber et al., 2015), this in turn may alter the observed gender gap. Indeed, when presented with open-ended questions, girls perform relatively better than boys, while boys obtain relatively better results when multiple choice questions are used; these results were found both in language and in mathematics (Bolger & Kellaghan, 1990; Lafontaine & Monseur, 2009; Lindberg et al., 2010; Reardon, Kalogrides, Fahle, Podolsky, & Zárate, 2018 – however, it is possible that this does not hold when item difficulty is high: see Beller & Gafni, 2000; Routitsky & Turner, 2003; Willingham & Cole, 1997).

Third, some evaluations create more *stress* in students due to their high-stakes, which may affect boys' and girls' academic performance differently. Exploiting a pure change in stakes in school examinations (high-, medium- and low-stakes), Azmat, Calsamiglia, and Iriberry (2016) found that girls performed worse in higher- compared to lower-stakes settings (in Catalan, but not in mathematics). Similarly, in settings combining high-stakes and competition, girls obtain lower results than expected across all subjects (Cai et al., 2018; Jurajda & Münich, 2011; Ors et al., 2013). One should note that it is uncertain in the literature to what extent the effect of stress is confounded with another factor, namely motivation to perform well at low-stakes tests. Indeed, there is some indication that girls are more motivated than boys to do their best at low-stakes tests compared to higher-stakes ones, which may affect the sex differences in performance across tests the same way than stress (DeMars et al., 2013; Eklöf, 2007; OECD, 2015; O'Neil et al., 2005 - note however that it is not clear from these papers that motivation does improve performance at the test).

Fourth, some evaluations, due to their non-blind nature, may generate a *grading bias* based on students' gender. Past studies have shown in a variety of countries that girls obtain higher results in non-anonymous evaluations compared to anonymous ones – thus suggesting that there is a grading bias in favor of girls (Breda & Ly, 2015; Falch & Naper, 2013; Lavy, 2008; Protivínský & Münich, 2018; Terrier, 2015; but see Lafontaine & Monseur, 2009a).

To what extent do each of these evaluation characteristics affect differences between boys and girls in academic performance? How much do sex differences in performance depend on these differences in evaluation characteristics? In the present paper, we aimed to explore these questions by studying differences between boys and girls in

measured academic performance. We compared boys' and girls' results at three different types of evaluations for the same set of students (teacher evaluations, national examinations, and standardized achievement tests) from a large, representative sample of middle school students in France.

5.1.3 Method

Sample

We used data from the DEPP Panel 2007, a large cohort study led by the Direction de l'Évaluation, de la Prospective et de la Performance (DEPP; French Ministry of Education) containing rich data on 34,986 French students from their first year of middle school in 2007 (grade 6, 11 years old) to their last year of middle school (grade 9, 14 years old) (Trosseille et al., 2013). The study was compulsory and approved by the National Council for Statistical Information (CNIS) (visa n°2008A061ED and 2011A082ED), ensuring public interest and conformity with ethical, statistical and confidentiality standards. The sample was randomly selected from an exhaustive sampling frame, ensuring representativeness by balancing available characteristics (region, public/private status of the school, urban unit, school establishment, age of entry in grade 6). The sample was constituted in such a way as to be representative of the population of French middle school students, with a slight over-representation of students in schools belonging to the *Réseau Ambition Réussite* (Success Ambition Network – schools in disadvantaged areas). The present study focusses on the grade 9 wave, when three different measures of achievement were reported: National examination grades, teacher grades, and standardized test scores. Our working sample includes students for whom results at the three tests in French and mathematics were available ($N = 23,451$). Thus, sample size was not determined based on expected effect sizes but on available data. With this sample size, we are able to detect a sex difference of size $d = 0.037$ with 80% power. 52% of the participants in our working sample were girls. Students' average age in grade 9 was equal to 14.09 years ($SD = 0.42$).

Measures

Academic achievement. Three different measures of academic achievement, in French and mathematics, were collected in grade 9:

National examination grades: At the end of grade 9 (in June), all French students have to take the written tests of a national examination, the *Diplôme national du Brevet* (DNB). The tests, lasting 2 hours each, are graded anonymously by teachers, and assess school knowledge acquired throughout the school year. They are composed of open-ended questions. In French, the exam is divided in three parts: open-ended questions on reading comprehension and grammar; text dictation; and essay. In mathematics, the exam is divided in three parts as well, including: numerical activities (open-ended questions in arithmetic, algebra and statistics); geometrical activities (open-ended questions in geometry); and problem solving (open-ended questions on a real-world problem). The DNB written examinations are the first official, nationwide examination that students take, and the grade they obtain constitutes about 40% of their final grade at the DNB (the remaining 60% coming from teacher grades in grade 9 in all subjects). Therefore, they are relatively high stakes for students.²⁰

Teacher grades: Teacher average grades include grades at in-class tests as well as homework grades throughout the year – thereby influenced by students’ self-discipline (both in terms of diligence with respect to school work and behavior in class). The grades count in the final DNB grade and are of great importance for selection into high school. Therefore, they also are relatively high stakes for students.

Standardized tests: For the purpose of the Panel 2007 cohort study, the DEPP administered standardized tests to students. The mathematics test included short open-ended and multiple choice questions testing students in logic, mental arithmetic, problem solving, units and time calculations, and geometry (45 items) (F. Aubret & Blanchard, 1992; Blanchard & Berger, 1994; OECD, 2011). In French, two tests were administered: a cloze test (blank-filling task) composed of three short texts with missing logical connectors, determiners or pronouns (20 items) (J. Aubret et al., 2006); and a reading comprehension test in limited time (12 min) composed of three short texts, each followed by five short open-ended questions (J. Aubret et al., 2006). The tests were administered in April to May. Internal consistency was good (see **Table A1**). The same tests were administered in grade 6 (except for some items that changed to match students’ level), with high correlations between both mathematics and cloze tests in grade 6 and grade 9, and a moderate

²⁰ We only had access to the final grade of the national examination, not to the different items composing the exam. Therefore, we could not compute reliability and consistency indices. We nevertheless provided the correlations between grades at the national examination, teacher grades and standardized examination as a proxy for reliability and validity (in Table A2).

correlation for the reading comprehension test, suggesting an acceptable test-retest reliability (see **Table A1**).

Note that in the standardized tests, both in mathematics and French, open-ended questions did not require students to write full sentences, but merely to write down a short answer. In contrast, both teacher evaluations and national examinations expected student to write full sentences as answers. **Table A2** shows correlation between the three types of evaluation, in Mathematics and in French. **Table 1** summarizes the main characteristics of the three types of evaluation available. All scores were standardized with zero mean and unit standard deviation.

Table 1

Characteristics of the three types of tests available in grade 9

Characteristics	Teacher grades	National examinations	Standardized tests
Anonymous scoring	No	Yes	Yes
Stakes	High	High	Low
Assessing writing skills	Yes	Yes	No
Role of self-discipline	Yes	No	No

Covariates. We used the following variables in our analysis in order to control for students' differences in socio-demographic and cognitive characteristics:

Socio-demographic variables: As part of the cohort study, parents had to fill in questionnaires in grade 9. We used information extracted from this questionnaire in our analysis in order to control for socio-demographic factors: parental education (average of the years of education of the mother and the father), household monthly income (in logarithmic scale), number of books and CDs at home, age of entry in grade 6, extracurricular activities, and being schooled in a priority school.

Cognitive variables: Along with standardized tests assessing academic performance, students completed a nonverbal intelligence test in grade 9, Chartier's Reasoning Test on Playing Cards (*Raisonnement sur Cartes de Chartier*, RCC)

(Cronbach's $\alpha=0.87$) (Terriot, 2014). They also filled in a questionnaire measuring their perceived self-efficacy in three different aspects (autoregulation, social, and academic), the Children's Perceived Self-Efficacy scales (Bandura, 1990), from which three factor scores were constructed using confirmatory factor analysis (Cronbach's α above 0.80 for each of the three scores). Lastly, they answered a questionnaire measuring school related motivation (intrinsic motivation, extrinsic motivation, and amotivation) derived from the Academic Self-Regulation Questionnaire (Ryan & Connell, 1989), from which three factor scores were constructed using confirmatory factor analysis (Cronbach's $\alpha=0.85$ for intrinsic motivation and 0.68 for extrinsic motivation and amotivation).

Analyses

Descriptive statistics. We first used descriptive statistics to understand how the gender gap in performance varies with the subject and evaluation type. We assessed the effect size of the difference for each evaluation score by computing Cohen's d , i.e. the standardized difference between the male and female students' means.

Difference-in-differences. In a second step, we used difference-in-differences regressions to assess the significance of the effect of evaluation type on the gender gap in achievement in French and mathematics. This method has already been used in sex differences studies, in particular to assess the existence of a grading bias (Breda & Ly, 2015; Lavy, 2008; Terrier, 2015). As is summarized in **Table 1**, we assume that the three evaluation types (national examinations, teacher evaluations, and standardized tests) have different characteristics. Independently from socio-demographic and individual factors which affect performance similarly across evaluations, results at standardized tests are supposed to reflect solely students' ability in the subject; results at the national examinations are, in addition, dependent on students' writing skills and their ability to cope with stress (induced by the high stakes involved)²¹; and results at teacher evaluations are, in addition to all of the above, influenced by students' self-discipline and teachers' biases (since they are not anonymous). It is likely that these characteristics affect male and female students differently. Based on the literature, we hypothesize for example that female students would have an advantage when the evaluation relies more heavily on writing skills,

²¹ As noted in the introduction, the effect of stress on high-stakes tests versus low-stakes tests may be confounded with motivation to perform at low-stakes tests, but it is not possible to disentangle them.

when it is influenced by self-discipline, and when it is not anonymous. Similarly, we hypothesize that they would be at a disadvantage when the stakes are higher (Azmat et al., 2016). In order to assess the relative influence of these different factors on the achievement gap between male and female students, we analyzed how the gender gap changes depending on the evaluation used. The presence of the three scores (standardized score, teacher grade, and national examinations grade) in each subject (French and mathematics) allowed us to use a difference-in-differences estimation strategy with three different conditions. The score obtained by a student i on evaluation j (*Score* – standardized with zero mean and unit standard deviation) depends on her sex (dummy variable *Fem* equal to 1 if the student is a female) and on the type of the evaluation (dummy variables *Teach* and *Std* equal to 1 for teacher evaluation or standardized test, respectively – the national examinations constitute the reference condition here). We may thus write the equation of the score for student i at evaluation j as follow:

$$\text{Score}_{ij} = \alpha + \gamma \text{Fem}_i + \beta_1 \text{Teach}_{ij} + \beta_2 \text{Std}_{ij} + \delta_1 (\text{Fem}_i \times \text{Teach}_{ij}) + \delta_2 (\text{Fem}_i \times \text{Std}_{ij}) + u_{ij} \quad (1)$$

Our coefficients of interest are those of the interaction terms, δ_1 and δ_2 . Estimating δ_1 and δ_2 with equation (1) is equivalent to estimating them with a difference equation where the difference in scores is the dependent variable (Lavy, 2008). Statistical analyses were carried out with the software SAS²². In order to obtain δ_1 and δ_2 , we estimated equation (1) using generalized estimating equations (Liang & Zeger, 1986) with standard errors clustered at the individual level in order to correct for dependence among our repeated observations (PROC GENMOD)²³. We also estimated equation (1) controlling for individual characteristics by including socio-demographic and cognitive variables as covariates. Missing data in covariates (see Table 2) were dealt with using multiple imputation (Rubin, 1987) (PROC MI and PROC MIANALYZE with MCMC method and 10 imputed datasets).

²² The script is available on the Open Science Framework with the link: https://osf.io/c37hb/?view_only=d227c87f5ffd439fbb5d8cc0ec84c5a7. Data are available on request on the Quêtelet PROGEDO French data archives for human and social sciences: <http://www.progedo-adisp.fr/enquetes/XML/lil-0955.xml>.

²³ We could not cluster at the school or class level because such information was not available.

δ_1 measures to what extent the discrepancy between teacher evaluation scores and national examinations (DNB) scores differs between male and female students. Similarly, δ_2 reflects to what extent the discrepancy between standardized test scores and national examinations (DNB) scores differs between male and female students. We can indeed derive from the above equation that:

$$(\text{Teach_Score} - \text{DNB_Score})^{\text{Fem}} - (\text{Teach_Score} - \text{DNB_Score})^{\text{Male}} = \delta_1 \quad (2a)$$

$$(\text{Std_Score} - \text{DNB_Score})^{\text{Fem}} - (\text{Std_Score} - \text{DNB_Score})^{\text{Male}} = \delta_2 \quad (2b)$$

We can further express the score at each type of evaluation – hence δ_1 and δ_2 – in terms of the evaluation's hypothesized characteristics. As stated above, standardized tests should solely measure ability in the subject; while both national examinations and teacher evaluation scores should reflect also their writing skills, and their ability to cope with stress. In addition, teacher evaluations may reflect potential gender biases in grading, and students' discipline. We can formalize this as follow:

$$\text{Std_Score}_i = \alpha_1 \text{ability}_i + \epsilon_i \quad (3a)$$

$$\text{DNB_Score}_i = \alpha_1 \text{ability}_i + \alpha_2 \text{writing}_i + \alpha_3 \text{stress}_i + \mu_i \quad (3b)$$

$$\text{Teach_Score}_i = \alpha_1 \text{ability}_i + \alpha_2 \text{writing}_i + \alpha_3 \text{stress}_i + \alpha_4 \text{discipline}_i + \alpha_5 \text{Fem}_i + \eta_i \quad (3c)$$

The variables *ability*, *writing*, *stress* and *discipline* represent, respectively, a student's general ability in the subject, her writing skills, her ability to cope with stress, and her self-discipline – characteristics that we do not measure directly. We assume that each of these unobserved factors affects results at the different evaluations the same way, for males and females (i.e. the coefficients α_1 , α_2 , α_3 and α_4 are assumed to be the same across evaluations and sexes for a given subject). The last coefficient α_5 represents the grading bias (such that if it is positive, there is a bias in favor of girls, and if it is negative, there is a bias in favor of boys). Injecting equations (3a), (3b) and (3c) into equations (2a) and (2b), we obtain:

$$\delta_1 = \alpha_4 (\text{discipline}^{\text{Fem}} - \text{discipline}^{\text{Male}}) + \alpha_5 \quad (4a)$$

$$\delta_2 = -\alpha_2 (\text{writing}^{\text{Fem}} - \text{writing}^{\text{Male}}) - \alpha_3 (\text{stress}^{\text{Fem}} - \text{stress}^{\text{Male}}) \quad (4b)$$

Obviously, it will not be possible to determine the values of all the unknowns, since the observed variables will be fewer. It will nevertheless be possible to make some valuable inferences. Indeed, δ_1 reflects the difference between male and female students that can be attributed to grading bias and discipline, while δ_2 reflects the difference between male and female students that can be attributed to writing skills and the ability to cope with stress. The signs of δ_1 and δ_2 , estimated with equation (1), can thus inform us on the relative effect of the different unobserved evaluation factors on the gender gap in performance. Since higher general ability, writing skills, ability to cope with stress and discipline are associated with higher test scores, α_1 , α_2 , α_3 and α_4 are assumed to be positive. Based on the extant literature, we expect female students to exhibit on average higher levels of self-discipline than their male counterparts. We also expect the grading bias to be in favor of female students (positive α_5). Therefore, we expect to find a positive δ_1 . Besides, we expect female students to have better writing skills, but to cope less well with stress. Thus, the two terms in equation (4b) are expected to be of opposite sign. A positive δ_2 would thus mean that the ability to cope with stress plays a more important role than writing skills in the gender gap in performance (whether it stems from the respective sizes of the α_2 and α_3 coefficients or the magnitude of the gender gaps in writing ability and stress). If δ_2 is null, then this means that both factors cancel out. Lastly, a negative δ_2 would suggest that writing skills play a more important role than the ability to cope with stress in the gender gap in performance.

5.1.4 Results

Sex differences between subjects and evaluation types

Table 2 shows descriptive statistics for all academic achievement variables as well as socio-demographic variables and cognitive variables by sex, and the effect size of the difference between male and female students.

Table 2

Descriptive statistics for achievement, socio-demographic, and cognitive variables, by sex

Variables	N	Girls (N=12180)		Boys (N=11271)		Difference		p
		M or %	S.D.	M or %	S.D.	d or O.R. [C.I]		
<i>Academic achievement variables</i>								
French DNB grade (z-score)	23,451	0.18	0.96	-0.21	1.00	0.40	[0.37; 0.42]	<0.001
French teacher grade (z-score)	23,451	0.20	0.97	-0.22	0.99	0.42	[0.40; 0.45]	<0.001
French standardized test (z-score)	23,451	0.09	0.99	-0.09	1.00	0.18	[0.15; 0.21]	<0.001
Mathematics DNB grade (z-score)	23,451	-0.05	0.97	0.05	1.03	-0.11	[-0.13; -0.08]	<0.001
Mathematics teacher grade (z-score)	23,451	0.03	1.00	-0.03	1.00	0.05	[0.03; 0.08]	<0.001
Mathematics standardized test (z-score)	23,451	-0.19	0.99	0.21	0.97	-0.40	[-0.43; -0.38]	<0.001
<i>Socio-demographic variables</i>								
Parental education (years)	22,104	12.13	3.39	12.38	3.43	-0.07	[-0.10; -0.08]	<0.001
Household monthly income (EUR)	13,412	3309	3783	3448	4070	-0.04	[-0.07; 0.00]	0.0408
Number of books and CDs in the household (score)	22,200	0.08	0.89	0.10	0.89	-0.02	[-0.05; 0.00]	0.0683
Age of entry in grade 6 (years)	23,451	11.08	0.40	11.09	0.00	-0.03	[-0.05; 0.00]	0.0491
Extracurricular activities (score)	21,775	0.27	0.27	0.28	0.27	-0.04	[-0.06; -0.01]	0.0049
Schooled in a disadvantaged area (% , OR)	21,661	17.50	-	16.48	-	1.08	[1.00; 1.15]	0.0379
<i>Cognitive variables</i>								
Non-verbal intelligence (out of 30)	23,306	18.85	5.59	18.82	6.02	0.00	[-0.02; 0.03]	0.7635
Perceived self-efficacy in autoregulation (z-score)	23,236	0.02	1.00	-0.02	1.00	0.04	[0.01; 0.07]	0.0023
Perceived academic self-efficacy (z-score)	22,517	0.10	0.97	-0.11	1.01	0.22	[0.19; 0.24]	<0.001
Perceived social self-efficacy (z-score)	22,692	-0.19	1.03	0.21	0.92	-0.41	[-0.44; -0.39]	<0.001
Intrinsic motivation (z-score)	23,089	0.04	0.99	-0.04	1.01	0.08	[0.057; 0.11]	<0.001
Extrinsic motivation (z-score)	23,089	-0.02	0.98	0.02	1.02	-0.04	[-0.06; -0.01]	0.0038
Amotivation (z-score)	23,089	-0.17	0.85	0.18	1.11	-0.36	[-0.38; -0.33]	<0.001

Fig. 1 displays the distribution of evaluation scores for male and female students by subject and evaluation type, and **Fig. 2** represents the effect sizes of the sex difference in performance across evaluations. Globally, female students performed better than their male peers in French, while the reverse was true in mathematics. However, their respective performance appears to depend on the evaluation used to evaluate them. There is indeed a clear trend as illustrated in **Fig. 1**, whereby both in French and mathematics female students scored the highest in teacher evaluations, followed by national examinations and lastly by standardized tests. Conversely, male students obtained their highest grades with standardized tests, followed by national examinations and teacher evaluations. Consequently, the gender gap in scores varies considerably across evaluation types, as illustrated in **Fig. 2**. This is most striking if we look at sex differences in mathematics scores across evaluation types: while the gap is clearly in favor of male students in the standardized test ($d = -0.41, p < 0.001$), it shrinks but is still in favor of male students in the national examination ($d = -0.11, p < 0.001$), and is almost null and reversed in the teacher examination ($d = 0.05, p < 0.001$). In French as well, the gap appears to vary across evaluation types, although it is in favor of female students in all three evaluations: the effect size is medium in the teacher evaluation ($d = 0.43, p < 0.001$) as well as in the national examination ($d = 0.40, p < 0.001$), but is small in the standardized test ($d = 0.18, p < 0.001$). In order to properly measure this apparent effect of evaluation type on the gender gap in achievement across subjects, we turn to the difference-in-differences estimation.

Figure 1. Distribution of test scores, by subject, sex and test type. The solid black lines represent group means and the beans the smoothed density curves. Source: MENESR DEPP.

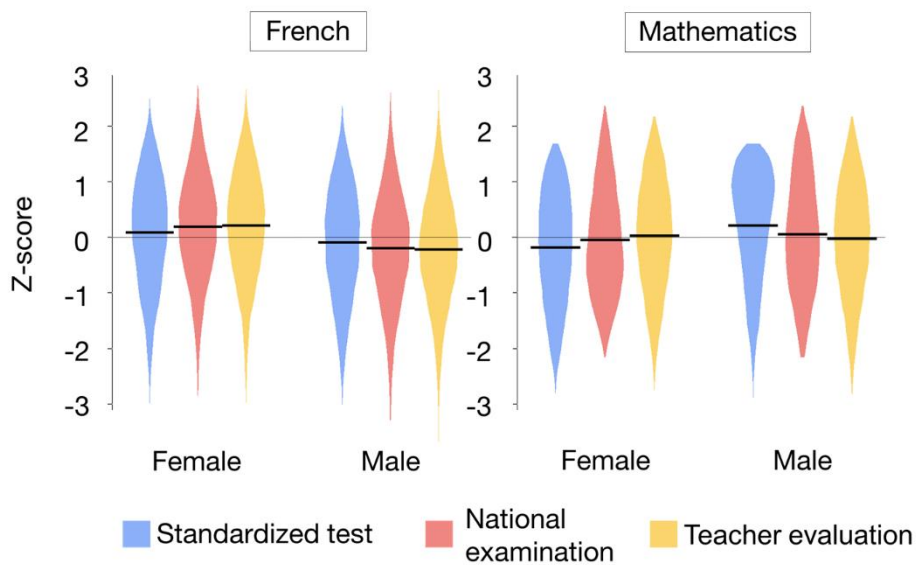
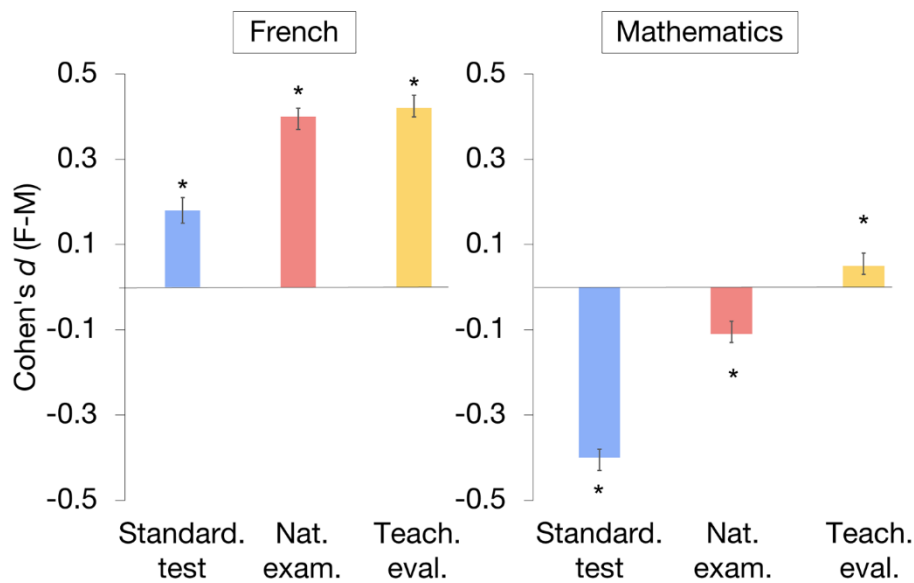


Figure 2. Difference in test scores between male and female students, by subject and test type. Error bars indicate 95% confidence intervals; and stars (*) indicate $p < 0.001$. “Standard. test” stands for standardized tests; “Nat. exam.” stands for national examinations; and “Teach. eval.” stands for teacher evaluations. Source: MENESR DEPP.



3.2. Relative effects of evaluation characteristics on sex differences

Table 3 displays results from the difference-in-differences regression without covariates. The coefficients of interest are the two interaction terms.

Table 3
Difference-in-differences regression estimates

Parameters	French		Mathematics	
	β [C.I.]	p	β [C.I.]	p
Female	0.40 [0.37; 0.42]	<0.001	-0.11 [-0.13; -0.08]	<0.001
Teacher evaluation	-0.01 [-0.03; 0.00]	0.0492	-0.08 [-0.10; -0.07]	<0.001
Standardized test	0.11 [0.10; 0.13]	<0.001	0.16 [0.14; 0.17]	<0.001
Female x Teacher evaluation (δ_1)	0.03 [0.01; 0.05]	0.0056	0.16 [0.14; 0.18]	<0.001
Female x Standardized test (δ_2)	-0.22 [-0.24; -0.20]	<0.001	-0.30 [-0.32; -0.28]	<0.001

Source: MENESR DEPP.

The coefficient δ_1 of the interaction term between *Female* and *Teacher evaluation* is significantly different from zero and positive, both in French and mathematics. Thus, female students increased their performance in teacher evaluations relative to the national examinations more than their male counterparts. According to our analysis of equation (4a), this may be due to greater self-discipline for homework and regular study in females, or teacher bias in favor of females, both factors likely to increase scores in teacher evaluations relative to national examinations. The effect is very small in French ($\delta_1 = 0.03$, $p = 0.0056$), but larger in mathematics ($\delta_1 = 0.16$, $p < 0.001$).

The second coefficient of interest, δ_2 , is that of the interaction term between *Female* and *Standardized test*. It is significantly different from zero, negative and of moderate size both in mathematics ($\delta_2 = -0.30$, $p < 0.001$) and in French ($\delta_2 = -0.22$, $p < 0.001$). Indeed, female students' performance was lower in standardized tests than at the national examinations, whereas that of their male peers was higher in standardized tests than at the national examinations. As indicated in equation (4b), the negative sign of δ_2 suggests that writing skills play a more important role in the observed gender gap in performance than the ability to cope with stress. Indeed, female students scored lower in the standardized tests despite the fact that these tests were lower stakes and therefore less stressful than the national examinations.

As can be seen from **Table A3**, controlling for students' socio-demographic and cognitive factors had little effect on the estimates of δ_1 and δ_2 . Hence these characteristics do not explain the gender gap in performance across evaluation types. As a last step, in order to ensure that the estimated effects do not reflect the effects of interactions between test type and other variables that would correlate with sex, we added interaction terms between our evaluation type dummies and all other covariates. The results are reported in **Table A4**. We can see that our coefficients of interest (the interactions between sex and evaluation type) are slightly reduced. In French, δ_1 is still non-significant and reduces from 0.03 in the regression without interactions, to -0.02 when adding in the interaction terms; while δ_2 reduces from -0.22 to -0.18 ($p < 0.001$). In mathematics, δ_1 reduces from 0.16 to 0.12 ($p < 0.001$) and δ_2 from -0.30 to -0.25 ($p < 0.001$). This decrease is due to the fact that other covariates significantly interacted with evaluation type, with a practically meaningful effect: being schooled in a disadvantaged area, and academic self-efficacy. However, these other interactions did not reduce by much the sex by evaluation type interaction terms, which thus strengthens our results.

5.1.5 Discussion

Evaluation conditions affect sex differences in achievement

This paper aimed at better understanding the relative influences of evaluation characteristics on sex differences in academic achievement. In line with previous findings, we found that girls performed better than boys in French, and worse in mathematics, when assessed with standardized achievement tests (Else-Quest et al., 2010; Hedges & Nowell, 1995; Hyde et al., 1990; Hyde & Linn, 1988; OECD, 2015; but see Lindberg et al., 2010), and that they performed better in both when assessed with teacher evaluations (Daniel Voyer & Voyer, 2014). Contrary to previous findings on national examinations by Deary et al. (2007), showing that British female students scored significantly higher than males in both English ($d = 0.41$) and mathematics ($d = 0.03$), we found that French female students performed higher in French ($d = 0.40$), but lower in mathematics ($d = -0.11$). This result may be due to socio-economic and cultural differences. For instance, the UK's Global Gender Gap Index (GGI), which measures gender parity in the areas of health, education, economy and politics, was higher than France's in the 2000s (Hausmann et al., 2010). Moreover, boys reported that their parents find mathematics more important to study and

for their future career than girls do, and this difference is lower in the UK (Stoet et al., 2016). Both the GGI and the gender gap in parental mathematics valuation are correlated with the gender gap in mathematics achievement in PISA (Stoet et al., 2016). Although these results contrast with the larger advantage of boys over girls in PISA mathematics test in the UK compared to France, it is possible that such socio-economic and cultural variations are more strongly associated with results at national examinations than at lower-stakes standardized achievement tests such as PISA.

Interpretations of the effect of evaluation conditions

The difference-in-differences models confirmed that girls over-performed in teacher evaluations and under-performed in standardized achievement tests, compared to national examinations. These effects were larger in mathematics than in French.

Females' greater relative performance in mathematics during teacher evaluations (compared to national examinations) may result from superior self-discipline in mathematics, a greater effect of self-discipline on mathematics, or from teachers having a greater grading bias in favor of females in mathematics. While our data does not allow us to distinguish these possibilities, the third option seems likely, as evidence for a grading bias in favor of girls has previously been reported in France in mathematics but not in French (Breda & Ly, 2015; Terrier, 2015).

Females' lower relative performance in mathematics on standardized tests (compared to national evaluations) may be due to several factors: a larger sex difference in writing skills in mathematics; a larger effect of writing skills on performance in mathematics; a smaller sex difference in stress in mathematics; or a lower effect of stress on performance in mathematics (or, undistinguishably, a lower motivation to perform well at low-stakes standardized tests in mathematics). While these possibilities may seem counter-intuitive, we are unable to adjudicate between them based on our data or on plausibility alone. Yet, Azmat, Calsamiglia, and Iriberry's (2016) findings that girls performed worse in higher-stakes settings in Catalan but not in mathematics suggest that there may be a smaller sex difference in stress in mathematics compared to French.

Limitations

In light of the following limitations, conclusions must be interpreted with caution. First, our inferences regarding the extent to which each characteristic affects sex differences

are purely based on assumptions from the literature, since we did not have the data to measure these characteristics. Further research with more comprehensive datasets are needed to address this issue, as it is also possible that we omitted other evaluation characteristics that may affect the gap across evaluations. For example, in mathematics, the presence of problems with spatially based solution strategies and multiple solution paths may influence the sex difference in favor of boys (Gallagher et al., 2002).

Second, our model is somewhat simplistic: it assumes that there are no interactions between the different evaluation characteristics, and that students' ability and evaluation characteristics affect results in the different evaluations in the same way across sexes. Although these assumptions were necessary in our model to draw simple interpretations regarding the effects of these factors on the gender gap, it is possible that they might not hold.

Another potential limitation is our assumption that male and female students' handwriting are indistinguishable in anonymous evaluations, which is why we only included a potential grading bias in teacher evaluations. However, Baird (1998) showed that grades are not affected by the gender style of handwriting and Breda and Ly (2015) found that the percentage of correct guesses of students' gender based on handwritten anonymous exam is only 68.6%. Furthermore, even if gender could be partly detected and induced a grading bias in anonymous examinations, this would only underestimate the grading bias in our study (and in previous studies on the subject).

Despite these limitations, our data clearly showed that there are differences in the gender gap in achievement across evaluations, which must depend on differences in evaluation characteristics.

Conclusions and practical implications

Sex differences in academic achievement is a hot topic for which multiple interpretations have been proposed. One possible interpretation is that such differences do not really exist in terms of sheer academic competence, but that they emerge only as a result of unfair evaluation conditions. Our results confirm that sex differences are modulated by multiple evaluation conditions. This multiplicity makes it difficult to attribute sex differences in a given evaluation or a given subject to a specific factor. For instance, in our study, it seems that stress due to perceived high-stakes only played a limited role in explaining the female disadvantage in mathematics, since girls actually succeeded better in

mathematics at the high stakes national examination than at the low stakes standardized tests – and this difference was smaller in French. Modifying these factors can alter sex differences substantially, sometimes even reversing them. However, this does not mean that sex differences are solely the product of evaluation conditions, and that they might entirely disappear under “ideal” testing conditions. In this regard, it is noteworthy that changing some factors may reduce the gap in one subject, while it may increase it in another subject. For example, we observed that using teacher evaluations reduced the gap in mathematics, but increased it in French. Therefore, finding testing conditions that globally reduce sex differences is complex. One may also question whether this is a worthy goal. Indeed, it is important to distinguish adverse impact from fairness (Halpern, 2002): finding significant sex differences in a given evaluation does not imply that the evaluation is unfair. A fair evaluation is one which is comparably valid for both sexes, meaning that it evaluates accurately the skills targeted for each group. To have a correct picture of students’ skills and progress, we should thus strive to eliminate factors that are not intended to be measured and that differentially affect female and male students. Hence, fair evaluations would in theory be low-stakes and corrected anonymously. In the context of our study, the evaluation that filled both of these conditions (the standardized test) was the one which reduced the gap the most in French, but increased it the most in mathematics. However, this test was based on multiple choice questions and short answers, in which boys tend to perform better. Ultimately, the response format does not affect fairness per se: the choice should depend on which response mode most accurately taps the target abilities. If anything, it would seem appropriate to balance different response formats as much as allowed by the test setting. In a test imbedded in cohort study as the standardized test that we used here, it would seem costly to introduce open-ended questions, as the time students can spend on the test is reduced. In the case of a national examination like the DNB which is already constituted of open-ended questions, introducing multiple choice items seems easy and reasonable.

5.1.6 Appendix

Table A1*Internal consistency for standardized tests scores*

Test	Cronbach's alpha	Correlation with the same score in grade 6
Mathematics	0.935	0.83
Cloze test	0.817	0.73
Reading comprehension	0.816	0.56

Table A2*Correlations between national examinations, standardized tests and teacher grades in French and Mathematics*

	National examination	Standardized test	Teacher grades
French			
National examination	1.00		
Standardized test	0.67	1.00	
Teacher grades	0.72	0.63	1.00
Mathematics			
National examinations	1.00		
Standardized tests	0.73	1.00	
Teacher grades	0.75	0.65	1.00

Table A3

Difference-in-differences regression estimates, controlling for students' characteristics

Parameters	French			Mathematics		
	β	[C.I.]	p	β	[C.I.]	p
Female	0.31	[0.28; 0.33]	<0.001	-0.19	[-0.21; -0.17]	<0.001
Teacher evaluation	-0.01	[-0.03; 0.01]	<0.001	-0.08	[-0.10; -0.06]	<0.001
Standardized test	0.11	[0.09; 0.13]	0.1868	0.16	[0.14; 0.17]	<0.001
Female x Teacher evaluation (δ_1)	0.03	[0.00; 0.06]	<0.001	0.16	[0.13; 0.19]	<0.001
Female x Standardized test (δ_2)	-0.22	[-0.24; -0.19]	0.0586	-0.30	[-0.33; -0.27]	<0.001
Non-verbal intelligence	0.05	[0.05; 0.05]	<0.001	0.07	[0.07; 0.07]	<0.001
Parental education	0.05	[0.05; 0.05]	<0.001	0.05	[0.05; 0.05]	<0.001
Household income	0.06	[0.04; 0.07]	<0.001	0.09	[0.07; 0.1]	<0.001
Books and CDs in the household	0.10	[0.09; 0.11]	<0.001	0.06	[0.05; 0.06]	<0.001
Extracurricular activities	0.16	[0.14; 0.19]	<0.001	0.11	[0.09; 0.13]	<0.001
Schooled in a disadvantaged area	-0.12	[-0.14; -0.1]	<0.001	-0.14	[-0.16; -0.13]	<0.001
Self-efficacy in autoregulation	0.09	[0.08; 0.10]	<0.001	0.03	[0.02; 0.04]	<0.001
Social self-efficacy	-0.13	[-0.14; -0.13]	<0.001	-0.14	[-0.15; -0.14]	<0.001
Academic self-efficacy	0.29	[0.28; 0.30]	<0.001	0.27	[0.26; 0.28]	<0.001
Intrinsic motivation	-0.05	[-0.06; -0.03]	<0.001	0.02	[0.01; 0.03]	<0.001
Extrinsic motivation	0.01	[0.00; 0.02]	<0.001	0.00	[-0.01; 0.01]	<0.001
Amotivation	0.00	[-0.01; 0.01]	0.0381	0.02	[0.01; 0.02]	0.5113

Source: MENESR DEPP

Table A4

Difference-in-differences regression estimates, controlling for students' characteristics and interactions between students' characteristics and evaluation type

Parameters	French		Mathematics	
	β [C.I.]	p	β [C.I.]	p
Female	0.31 [0.29; 0.33]	<.0001	-0.20 [-0.22; -0.18]	<.0001
Teacher evaluation	0.31 [0.06; 0.55]	0.0145	0.21 [-0.03; 0.45]	0.0876
Standardized test	-0.17 [-0.42; 0.07]	0.1553	-0.17 [-0.39; 0.04]	0.1165
Female x Teacher evaluation (δ_1)	-0.02 [-0.05; 0.01]	0.2378	0.12 [0.09; 0.15]	<.0001
Female x Standardized test (δ_2)	-0.18 [-0.21; -0.15]	<.0001	-0.25 [-0.28; -0.22]	<.0001
Parental education x Teacher evaluation	-0.01 [-0.01; 0]	0.0012	-0.01 [-0.02; -0.01]	<.0001
Parental education x Standardized test	0.00 [-0.01; 0]	0.2016	-0.01 [-0.02; -0.01]	<.0001
Household income x Teacher evaluation	-0.02 [-0.05; 0.02]	0.3085	-0.01 [-0.04; 0.02]	0.5715
Household income x Standardized test	-0.02 [-0.05; 0.02]	0.3232	-0.01 [-0.04; 0.02]	0.3979
Books and CDs x Teacher evaluation	-0.02 [-0.04; 0]	0.0516	-0.03 [-0.05; -0.01]	0.0013
Books and CDs x Standardized test	0.04 [0.02; 0.06]	<.0001	0.02 [0; 0.04]	0.0242
Extracurricular activities x Teacher evaluation	-0.01 [-0.06; 0.05]	0.8535	0.01 [-0.05; 0.06]	0.8210
Extracurricular activities x Standardized test	-0.01 [-0.06; 0.05]	0.7723	-0.02 [-0.07; 0.04]	0.5127
Schooled in a disadvantaged area x Teacher evaluation	0.19 [0.15; 0.23]	<.0001	0.28 [0.24; 0.32]	<.0001
Schooled in a disadvantaged area x Standardized test	-0.08 [-0.12; -0.04]	<.0001	0.01 [-0.02; 0.05]	0.4849
Non-verbal intelligence x Teacher evaluation	0.00 [0; 0]	0.8188	0.00 [0; 0]	0.1595
Non-verbal intelligence x Standardized test	0.02 [0.02; 0.02]	<.0001	0.03 [0.02; 0.03]	<.0001
Self-efficacy in autoregulation x Teacher evaluation	-0.03 [-0.06; -0.01]	0.0032	-0.02 [-0.04; 0]	0.0536

Self-efficacy in autoregulation x Standardized test	0.07	[0.04; 0.09]	<.0001	0.03	[0.01; 0.05]	0.0039
Social self-efficacy x Teacher evaluation	-0.05	[-0.07; -0.03]	<.0001	-0.03	[-0.05; -0.02]	0.0002
Social self-efficacy x Standardized test	0.04	[0.02; 0.06]	<.0001	0.07	[0.06; 0.09]	<.0001
Academic self-efficacy x Teacher evaluation	0.13	[0.1; 0.15]	<.0001	0.09	[0.07; 0.11]	<.0001
Academic self-efficacy x Standardized test	-0.08	[-0.1; -0.06]	<.0001	-0.11	[-0.13; -0.09]	<.0001
Intrinsic motivation x Teacher evaluation	0.03	[0; 0.06]	0.0211	0.02	[0; 0.05]	0.0777
Intrinsic motivation x Standardized test	-0.02	[-0.05; 0.01]	0.1142	-0.05	[-0.08; -0.03]	<.0001
Extrinsic motivation x Teacher evaluation	0.02	[-0.01; 0.04]	0.1783	0.02	[0; 0.04]	0.1221
Extrinsic motivation x Standardized test	-0.03	[-0.06; -0.01]	0.0083	0.00	[-0.02; 0.02]	0.8370
Amotivation x Teacher evaluation	0.02	[-0.01; 0.04]	0.1450	-0.01	[-0.03; 0.01]	0.5485
Amotivation x Standardized test	0.02	[0; 0.04]	0.0729	0.00	[-0.03; 0.02]	0.6708
Parental education	0.06	[0.05; 0.06]	<.0001	0.06	[0.05; 0.06]	<.0001
Household income	0.07	[0.05; 0.09]	<.0001	0.09	[0.07; 0.12]	<.0001
Books and CDs in the household	0.09	[0.08; 0.11]	<.0001	0.06	[0.05; 0.07]	<.0001
Extracurricular activities	0.17	[0.13; 0.21]	<.0001	0.11	[0.07; 0.15]	<.0001
Schooled in a disadvantaged area	-0.15	[-0.18; -0.13]	<.0001	-0.24	[-0.27; -0.21]	<.0001
Non-verbal intelligence	0.04	[0.04; 0.05]	<.0001	0.06	[0.06; 0.06]	<.0001
Self-efficacy in autoregulation	0.08	[0.06; 0.09]	<.0001	0.03	[0.01; 0.04]	0.0010
Social self-efficacy	-0.13	[-0.14; -0.12]	<.0001	-0.16	[-0.17; -0.15]	<.0001
Academic self-efficacy	0.28	[0.26; 0.29]	<.0001	0.28	[0.26; 0.29]	<.0001
Intrinsic motivation	-0.05	[-0.07; -0.03]	<.0001	0.03	[0.01; 0.05]	0.0010
Extrinsic motivation	0.02	[0; 0.03]	0.0707	-0.01	[-0.02; 0.01]	0.3072
Amotivation	-0.02	[-0.03; 0]	0.0564	0.02	[0; 0.03]	0.0135

Chapter 6 - General discussion

Learning to read, write and count is fundamental to unlock one's potential and flourish in modern society. These essential skills are acquired throughout schooling, building on a rich set of cognitive and socio-emotional skills, and influenced by a myriad of environmental, biological and genetic factors. This dissertation aimed at providing a greater understanding of what fosters or hampers the acquisition of academic skills in France, taking into account this multidimensionality. To do so, we have conducted a series of studies using longitudinal data from two French cohorts, the EDEN cohort and the DEPP Panel 2007, assessing the relative influences of a wide variety of factors on diverse aspects of academic achievement. In Chapter 2, we studied the extent of the association between intelligence and academic skills in France, firstly by assessing the strength of this relationship, as well as the socio-economic and cognitive predictors of academic skills and their progression beyond the role of IQ; secondly by investigating academic achievement among high-IQ students. In Chapter 3, we examined the preschool predictors of numeracy, assessing the relative predictive power of cognitive, socio-emotional and environmental factors on arithmetic skills as well as their mediation relationships; and we investigated the differential cognitive predictors of addition, subtraction and multiplication. In Chapter 4, we similarly studied the preschool cognitive, socio-emotional and environmental influences on the acquisition of different literacy skills, and their mediating relationships. Lastly, in Chapter 5, we examined sex differences in literacy and numeracy, assessing the influence of evaluation characteristics on these gaps. In this concluding chapter, we bring together the results from these different original contributions, synthesizing and discussing what we have learnt over the course of this dissertation about the cognitive, socio-emotional (Part 6.1), environmental and individual (Part 6.2) predictors of academic achievement in France; and acknowledging the general limitations inherent to our work (Part 6.3). Lastly, we discuss the practical implications of these findings, and conclude (Part 6.4).

6.1 Cognitive and socio-emotional foundations of academic skills

Our results corroborate the role of cognitive abilities as fundamental building blocks of academic skills, bringing new knowledge on their relative influences (6.1.1). Furthermore, we unveiled the influence of some socio-emotional abilities on academic achievement beyond cognitive skills (6.1.2).

6.1.1 Cognitive skills

This dissertation shed light on the specific influence (or lack thereof) of cognitive abilities on academic achievement, investigating general intelligence, language, visuospatial and fine motor skills.

General intelligence

While the tight link between intelligence and academic achievement had been amply evidenced in the past (Deary et al., 2007; Rohde & Thompson, 2007; Roth et al., 2015), several questions remained. First, the strength of this association had not been investigated in a large representative sample in the French context. Second, the role of general intelligence relative to other socio-emotional and environmental factors was unclear, since few studies included these factors simultaneously; and its influence on academic progression was similarly unknown. Third, evidence from unbiased, representative samples regarding the association between intellectual giftedness and academic achievement was lacking, as well as an estimation of the IQ-achievement relationship in the high-IQ range.

We investigated these questions in Chapter 2. In the large, representative sample of French middle school students that is the DEPP Panel 2007, we found that IQ was by far the most predictive factor of achievement at ages 11 and 14, far ahead of other environmental and socio-emotional factors (its standardized regression coefficient, about 0.5, was roughly double the size of the next most important predictor's, self-efficacy). However, its influence on academic progression throughout middle school was very limited (Chapter 2, Part 2.1). This positive association between intelligence and academic performance remained virtually the same in the high-IQ range, indicating that within

Chapter 6 - General discussion

intellectually gifted children, higher intelligence was associated with higher performance in a similar fashion as for the general population. Lastly, and contrary to popular belief on intellectually gifted children, we found in this large representative sample that high-IQ children had on average better academic results than their peers (Chapter 2, Part 2.2).

Going beyond general cognitive ability and academic achievement, we took advantage of the rich set of cognitive measures in the EDEN cohort to cast light on the relative roles of more specific cognitive factors – namely, language, visuospatial and fine motor skills – on the acquisition of separate academic skills (numeracy and literacy) in France.

Language abilities

Our results confirmed the foundational role of language abilities in the acquisition of academic skills, influencing literacy but also numeracy. Children with better language abilities in preschool performed better in reading and spelling at 11.5 years old, the strongest effect being on reading comprehension (standardized β equal to 0.5). In particular, phonological processing and verbal short-term memory were the best predictors of reading accuracy, while verbal short-term memory and language comprehension/conceptual knowledge best predicted spelling (Chapter 4). Children with better language skills at 5.5 also performed better when solving arithmetic word problems at 11.5, albeit to a lower extent (standardized β of 0.2) (Chapter 3, Part 3.1). However, not all types of arithmetic operations similarly relied on language skills: indeed, only multiplication at 11.5 was significantly predicted by preschool language, but it was not the case for addition nor for subtraction (Chapter 3, Part 3.2).

Visuospatial abilities

By contrast, the influence of early visuospatial abilities seemed to be more specific, supporting only numeracy but not literacy. Indeed, preschool visuospatial skills were not significantly associated with later literacy once early language was accounted for (Chapter 4). Therefore, visuospatial abilities do not seem to capture unique variance in literacy. However, they were the best predictor of arithmetic skills at 11.5 (standardized β of 0.3) (Chapter 3, Part 3.1) – yet, only addition and subtraction, but not multiplication, appeared to be supported by stronger visuospatial skills (Chapter 3, Part 3.2).

Fine motor abilities

Lastly, preschool fine motor skills did not appear to have any influence on later academic skills. Neither literacy nor numeracy abilities at 11.5 were significantly explained by fine motor skills at 5.5 once language and visuospatial skills were controlled for (Chapter 3, Part 3.1; Chapter 4). Thus, the association between fine motor abilities and academic skills seems to be fully accounted for by the overlap between fine motor and cognitive abilities, with no unique influence of fine motor skills on later academic performance.

6.1.2 Socio-emotional skills

Our work unveiled the specific roles played by some socio-emotional abilities in shaping academic skills, beyond cognitive skills.

Motivation and beliefs

Analyzing data from the DEPP Panel 2007, we studied the influence of conative abilities, assessing motivation and beliefs. Our results indicate that perceived self-efficacy was significantly associated with academic achievement at ages 11 and 14 – in fact, it was the most predictive factor after IQ (standardized β of 0.2-0.3). However, it did not influence academic progression between these two time points (Chapter 2, Part 2.1). In line with these results, we found that high-IQ children, who have higher academic skills than their peers, also had higher motivation, academic self-efficacy and self-regulation (Chapter 2, Part 2.2).

Internalizing and externalizing problems

Using the EDEN cohort, we were able to examine the relative influence of early socio-emotional problems on later literacy and numeracy. Two particular kinds of difficulties stood out. First, emotional symptoms displayed at 5.5 years old were marginally (not significant after correction) negatively associated with arithmetic (standardized β of -0.1; Chapter 3, Part 3.1) and literacy outcomes (standardized β of -0.1 for reading accuracy to -0.3 for reading comprehension; Chapter 4) at 11.5, after accounting for early cognitive abilities. Second, early inattention-hyperactivity symptoms were negatively associated with later literacy performance at 11.5 – in particular, reading accuracy and, marginally, spelling

(standardized β s of -0.2; Chapter 4). Thus, anxiety problems seem to affect the learning process in general, while hyperactivity-inattention issues seem to have a more specific role in reading acquisition.

6.2 Environmental and individual influences on academic achievement

In addition to cognitive and socio-emotional factors, this research investigated the pre- and post-natal environmental and individual determinants of academic skills, assessing the relative influences of parental socio-economic and cultural factors (6.2.1), sex (6.2.2), and birth and pre-natal factors (6.2.3).

6.2.1 Parental socio-economic and cultural factors

Our work confirmed the strong influence of parental education and related socio-economic factors on academic achievement, investigated in both samples.

Parental education

Parental education was one of the most important predictors of academic skills in all our studies, with a consistent effect of 0.2-0.3 standard deviation, which is in line with meta-analytic results (Sirin, 2005). Thus, the standardized β s of parental education were equal to 0.2 for general achievement (Chapter 2, Part 2.1), arithmetic skills (Chapter 3, Part 3.1), reading speed and spelling (although not significant), and to 0.3 for reading accuracy and comprehension (Chapter 4). However, this influence was reduced once past achievement was controlled for (standardized β of 0.05; Chapter 2, Part 2.1). Therefore, children with higher-educated parents performed better than their peers, and progressed more, although the latter effect was much smaller. Interestingly, this positive association between parental education and achievement was moderated by intellectual giftedness status, such that high-IQ children's performance was less related to their parents'

backgrounds than their peers', suggesting that a high IQ buffers against the effect of low SES on achievement – however, this effect was without practical significance.

Investigating the extent to which the effect of parental education on academic achievement was direct or mediated through early cognitive skills, we found a significant indirect effect on both literacy and numeracy, in particular through early language skills (standardized β s of 0.1 for arithmetic, reading accuracy and spelling), with large but non-significant direct effects (standardized β s of 0.1 for arithmetic and 0.2 for literacy skills) (Chapter 3, Part 3.1; Chapter 4). While further studies with more participants are required to conclude, these results suggest that parental education influences a child's academic performance at several stages, with early effects on cognitive skills that later affect the acquisition of academic skills, and later direct effects on achievement, for instance in the form of help with homework or material and cultural resources which enhance academic achievement.

Parental resources and involvement

Through the EDEN and DEPP Panel 2007 data, we were able to study the influence of a variety of parental socio-economic factors on academic skills beyond parental education, including material and cultural resources, as well as parental involvement with the child (e.g. early cognitive stimulation, help with homework, participation in school).

The school environment provided to the child, which was assessed through the enrolment status in a school located in a disadvantaged area (priority education school), was the most important environmental predictor of general academic achievement beyond parental education: students schooled in RAR and RRS schools had lower scores than their peers (β of -0.2), and progressed less during middle school (β of -0.07). By contrast, cultural objects in the house (books and CDs) had a positive influence on general academic achievement (β of 0.1), and a very small effect on progression during middle school (β of 0.02) (Chapter 2, Part 2.1). Household income was also positively associated with general academic skills in the DEPP Panel 2007, although to a lower extent (β of 0.1), and with academic skills gains (β of 0.03) (Chapter 2, Part 2.1). However, its influence on arithmetic and literacy, measured in EDEN, was null (Chapter 3, Part 3.1; Chapter 4). This discrepancy may be the result of sample differences, since the DEPP data was representative of the French population, while the EDEN sample at 11.5 was biased towards high-SES families. The provision of extracurricular activities was also marginally

Chapter 6 - General discussion

associated with general achievement (β of 0.05) but did not affect progression (Chapter 2, Part 2.1).

Parental involvement with the child, in the form of early cognitive stimulation (measured by the HOME questionnaire in the EDEN cohort) had a positive influence on later literacy (β of 0.2 on reading accuracy) (Chapter 4), but a negligible effect on arithmetic (Chapter 3, Part 3.1). Its effect on reading accuracy seemed to be mostly mediated through early cognitive skills, although no indirect nor direct effect appeared significant (Chapter 4). When assessing the influence of later parental involvement during middle school in the DEPP Panel 2007, we found mixed results. The occurrence of conversations between parents and child about school, and parental involvement in school life were positively related with academic achievement and academic progression but with very small effects (β s less than 0.03). Conversely, conversations with the child about their professional future, help provided to the child, and meetings with teachers were negatively associated with general academic achievement, and negatively influenced academic progression (even though their effect on progression was small or non-significant – standardized β s lower than 0.1). It is likely that unmeasured factors that are negatively associated with both academic achievement and with these factors (e.g., child behavior in class) partially explained this result, as it is hard to imagine that such factors intrinsically have a negative influence on academic skills.

6.2.2 Sex

Our research exposed sizeable differences between boys and girls in academic achievement in France – however, with important variations across skills. In Chapter 2, Part 2.1, we showed that girls had lower general academic achievement than boys (β of -0.1), with no differences in progression during middle school. When looking separately at the numeracy and literacy components of achievement in the EDEN cohort, we found a large difference in arithmetic skills (β of -0.5). Interestingly, this effect was mostly direct, i.e., unexplained by early differences in cognitive abilities (Chapter 3, Part 3.1). By contrast, there was no difference in any literacy skill (Chapter 4). When digging further into the potential causes that may explain discrepancies in sex differences across tests using the DEPP data, we found that, while overall girls performed better in literacy, and worse in

mathematics, this difference is sometimes heavily moderated by the evaluation type (Chapter 5). In the DEPP standardized test, the effect size in mathematics was large, comparable to the effect found in the EDEN cohort ($d = -0.41$). However, it was much smaller in the national examination ($d = -0.11$), and almost null and reversed in the teacher examination ($d = 0.05$). In literacy, the gap was small in the standardized test ($d = 0.18$), and large in the teacher evaluation ($d = 0.43$) and the national examination ($d = 0.40$). Given the different characteristics of each evaluation as highlighted in Chapter 5, these results are in line with the reported higher self-discipline of female students (Duckworth et al., 2015; Duckworth & Seligman, 2006; Kling et al., 2013) the grading bias in their favor (Breda & Ly, 2015; Falch & Naper, 2013; Lavy, 2008; Protivínský & München, 2018; Terrier, 2015) – making girls perform better in teacher evaluations than national examinations in both literacy and numeracy. In addition, the finding that girls perform better in national examinations than in standardized tests in both skills suggest that writing skills play a more important role than the ability to cope with stress in the gender gap in performance.

6.2.3 Pre-natal and birth factors

Lastly, data from the EDEN cohort allowed us to test the effect of pre-natal and birth factors on academic skills: number of siblings, gestational age, birthweight, age of mother, mother's alcohol and tobacco consumption during pregnancy, breastfeeding and maternal depression. None of these factors were significantly associated with arithmetic skills (Chapter 3, Part 3.1), nor to literacy skills (Chapter 4) once parental education and sex were controlled.

6.3 General limitations

In addition to the limitations already discussed in the different chapters, two main general considerations need to be re-emphasized: the lack of genetic factors in our study, which may account for part of the observed effects (6.3.1), and, relatedly, the correlational nature of our studies, preventing any causal interpretation (6.3.2).

6.3.1 Genetic confounding

Previous research has demonstrated the high heritability of academic skills. Analyzing data from 7,500 pairs of UK twins assessed at 7, 9 and 12 years old, Kovas et al. (2013) have shown that the heritability of literacy and numeracy skills was equal to 68% - higher than that of intelligence (42%). Smaller twin studies in Australia, the US, the Netherlands and Scandinavia have yielded similar estimates, ranging from 60% to 77% (Bartels et al., 2002; Byrne et al., 2009; Calvin et al., 2012; Olson et al., 2011). More recently, GWA studies have investigated genetic influences on educational attainment, identifying more than 1,000 genetic loci associated with it, with the resulting polygenic score explaining about 12% of its variance (Lee et al., 2018) and 15% of the variance in academic achievement at 16 years-old (Allegrini et al., 2019).

A significant part of this genetic influence on academic skills may be intertwined with the influence of environmental factors and other cognitive and socio-emotional skills. In Chapter 1, we explained how genetic factors play an important role in determining cognitive and socio-emotional abilities, and how environmental and genetic factors may be confounded, through the existence of gene-environment correlations. This phenomenon may also occur with academic skills. Let us illustrate the different forms that gene-environment correlations can take when applied to acquisition of academic skills. The first mechanism, passive gene-environment correlations, arises when parents with genes granting higher cognitive abilities also have higher levels of education and are thus able to provide an environment more likely to foster a child's academic success: hence, their child inherits these genes, and, in addition, grows up in a richer environment (better material and cultural resources, higher parental involvement) – and, conversely, children who evolve in disadvantaged environments are also more likely to have inherited genes for lower cognitive abilities. Illustrating this idea, a recent study by Cheesman et al. (2020) showed that polygenic scores are twice as predictive of educational attainment for non-adopted individuals than for adopted individuals (who grew up in environments less correlated with their genetic makeup). The second mechanism, evocative gene-environment correlations, occurs when children who get better grades (which is partly genetically determined) may be placed in better classes at school (e.g., with high-achieving or motivated peers) and given more opportunities to develop their abilities – and, on the contrary, children with difficulties may elicit fewer learning opportunities or be placed in school environments less conducive to academic success (see Harden et al., 2020, for the association between polygenic scores,

SES, tracking and achievement). Lastly, active gene-environment correlations arise because children with a higher genetic propensity to succeed at school are more likely to actively seek environments, peers and activities which foster academic skills (e.g. extracurricular activities, reading).

Consequently, there are significant genetic correlations between various measures of a child's environment, cognitive and socio-emotional abilities, and academic achievement. For instance, the correlation between perceived school environment, home environment, intelligence and achievement is partly explained by genetic factors (Krapohl et al., 2014); in fact, half of the correlation between achievement scores and family SES appears to be mediated genetically (Krapohl & Plomin, 2016). A moderate to large genetic overlap was also found between language abilities and reading fluency and reading comprehension, respectively (Tosto et al., 2017). As another example, 60% of the association between self-efficacy and achievement also appear to be accounted for by genetic factors (Greven et al., 2009; Krapohl et al., 2014). More recently, Malanchini et al. (2019) showed that differences in cognitive abilities, self-regulation, personality, motivation and attitudes towards learning fully explained the genetic variance in children's literacy and numeracy performance.

Therefore, the effects of cognitive, socio-emotional and environmental factors on academic skills assessed in the studies included in this dissertation are likely to be substantially mediated through genetic factors, and their net effects overestimated. The inclusion of polygenic scores capturing genetic influences on academic skills will enable future research to control for genetic differences between children and thus accurately estimate the net effects of such predictors on academic achievement. This may become possible in the near future in the EDEN cohort, with the collection of genetic data in participants.

6.3.2 Correlation and causality

It logically follows from the above limitation that our designs were not sufficient to infer causality. In addition to genes, several unmeasured variables that we were not able to include may have induced biases in our estimates, preventing us from deducing any causal effect from our various factors on academic skills – thus the denomination of these factors

as predictors. For example, we were not able to account for peer effects in school or class. Being in a class with more diligent and better achieving peers has been shown to increase one's academic results (Burke & Sass, 2013; Hanushek et al., 2003; McEwan, 2003; Zimmerman, 2003). Since a child's peers' level is also likely to be associated with the child's baseline academic skills, but also with his/her socio-economic background, not accounting for peer effects might have overestimated the influence of parental education for instance.

However, such correlational, observational studies remain extremely valuable for several reasons. Identifying the predictors of academic skills allows to detect children who are at-risk of future learning difficulties, and thus support them better. Besides, longitudinal studies inform future intervention studies which could properly test the causal effect of factors which have proven of interest (when ethically possible) – see, for instance, Hulme et al., (2012).

6.4 Practical implications and conclusion

To conclude, this research project allowed us to gain an enhanced understanding of the set of factors which contribute to individual differences in the acquisition of academic skills. These results have practical implications for practitioners and actors in the education sphere, and have the potential to inform future intervention studies. For instance, results from Chapter 2 directly talk to parents and teachers, correcting the beliefs that intellectually gifted children may be at risk of failure; Chapters 3 and 4 suggest that interventions promoting early language and targeting early emotional symptoms may enhance future literacy and mathematics abilities, while training early visuospatial skills may improve mathematics, and addressing early hyperactivity/inattention problems may prevent reading difficulties. Future research using randomized controlled trials in school or clinical settings are needed to assess these propositions. Besides, it follows from Chapter 3 that interventions aiming at reducing the gender gap in mathematics should focus on primary school. Future research should study more closely the evolution of the gender gap in mathematics during the primary school years in France to better understand how and when it develops. Lastly, Chapter 5 provides a nuanced vision of sex differences in literacy and

numeracy, emphasizing the sizeable effects of evaluation type and characteristics. Such effects need to be studied more thoroughly and be taken into account by examination and survey designers.

References

- Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-Analysis of Neurobehavioral Outcomes in Very Preterm and/or Very Low Birth Weight Children. *PEDIATRICS*, *124*(2), 717-728.
<https://doi.org/10.1542/peds.2008-2816>
- Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry*, *24*(6), 819-827. <https://doi.org/10.1038/s41380-019-0394-4>
- Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133-137. <https://doi.org/10.1016/j.lindif.2010.09.013>
- Ancker, J. S., & Kaufman, D. (2007). Rethinking Health Numeracy : A Multidisciplinary Literature Review. *Journal of the American Medical Informatics Association*, *14*(6), 713-721. <https://doi.org/10.1197/jamia.M2464>
- Andersson, Ulf. (2008). Working memory as a predictor of written arithmetical skills in children : The importance of central executive functions. *British Journal of Educational Psychology*, *78*(2), 181-203.
<https://doi.org/10.1348/000709907X209854>
- Ashcraft, M. H. (1992). Cognitive arithmetic : A review of data and theory. *Cognition*, *44*(1-2), 75-106. [https://doi.org/10.1016/0010-0277\(92\)90051-I](https://doi.org/10.1016/0010-0277(92)90051-I)
- Ashcraft, M. H. (2002). Math Anxiety : Personal, Educational, and Cognitive Consequences. *Current Directions in Psychological Science*, *11*(5), 181-185.
<https://doi.org/10.1111/1467-8721.00196>
- Aubret, F., & Blanchard, S. (1992). Valeur prédictive du test analytique de mathématique. *L'orientation scolaire et professionnelle*, *21*(4), 449-454.
- Aubret, J., Blanchard, S., & Sontag, J.-C. (2006). Évaluer les compétences des collégiens en 6e/5e. *L'orientation scolaire et professionnelle*, *35*(3), 446-473.
- Azmat, G., Calsamiglia, C., & Iriberry, N. (2016). Gender Differences in Response to Big Stakes. *Journal of the European Economic Association*, *14*(6), 1372-1400.
<https://doi.org/10.1111/jeea.12180>
- Bacete, F.-J. G., & Remírez, J. R. (2001). Family and Personal Correlates of Academic Achievement. *Psychological Reports*, *88*(2), 533-547.
<https://doi.org/10.2466/pr0.2001.88.2.533>
- Baird, J. (1998). What's in a name? Experiments with blind marking in A-level examinations. *Educational Research*, *40*(2), 191-202.
<https://doi.org/10.1080/0013188980400207>

- Banach, R., Boskovic, R., Einarson, T., & Koren, G. (2010). Long-Term Developmental Outcome of Children of Women with Epilepsy, Unexposed or Exposed Prenatally to Antiepileptic Drugs : A Meta-Analysis of Cohort Studies. *Drug Safety*, 33(1), 73-79. <https://doi.org/10.2165/11317640-000000000-00000>
- Bandura, A. (1990). *Multidimensional Scales of Perceived Self-Efficacy*. Stanford University.
- Barre, N., Morgan, A., Doyle, L. W., & Anderson, P. J. (2011). Language Abilities in Children Who Were Very Preterm and/or Very Low Birth Weight : A Meta-Analysis. *The Journal of Pediatrics*, 158(5), 766-774.e1. <https://doi.org/10.1016/j.jpeds.2010.10.032>
- Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions : Can we really discard any counting-based account? *Cognition*, 128(1), 35-44. <https://doi.org/10.1016/j.cognition.2013.02.018>
- Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Genetic and Environmental Influences on the Development of Intelligence. *Behavior Genetics*, 32(4), 237-249. <https://doi.org/10.1023/A:1019772628912>
- Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and Later Mortality Risk : Systematic Review. *Annals of Epidemiology*, 17(4), 278-288. <https://doi.org/10.1016/j.annepidem.2006.07.010>
- Beauregard, J. L., Drews-Botsch, C., Sales, J. M., Flanders, W. D., & Kramer, M. R. (2018). Preterm Birth, Poverty, and Cognitive Development. *Pediatrics*, 141(1), e20170509. <https://doi.org/10.1542/peds.2017-0509>
- Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, 107(5), 1860-1863. <https://doi.org/10.1073/pnas.0910967107>
- Beller, M., & Gafni, N. (2000). Can Item Format (Multiple Choice vs. Open-Ended) Account for Gender Differences in Mathematics Achievement? *Sex Roles*, 42(1-2), 1-21. <https://doi.org/10.1023/A:1007051109754>
- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care? *Child Development*, 78(2), 681-701. <https://doi.org/10.1111/j.1467-8624.2007.01021.x>
- Ben Ali, L., & Voure'h, R. (2015). Evolution des acquis cognitifs au collège au regard de l'environnement de l'élève. Constat et mise en perspective longitudinale. *Éducation & formations*, 86-87.
- Bénabou, R., Kramarz, F., & Prost, C. (2009). The French zones d'éducation prioritaire : Much ado about nothing? *Economics of Education Review*, 28(3), 345-356. <https://doi.org/10.1016/j.econedurev.2008.04.005>

- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low Health Literacy and Health Outcomes : An Updated Systematic Review. *Annals of Internal Medicine*, *155*(2), 97. <https://doi.org/10.7326/0003-4819-155-2-201107190-00005>
- Binet, A., & Simon, T. (1916). *The development of intelligence in children (The Binet-Simon Scale)*. Williams & Wilkins Co.
- Binet, A., & Simon, Th. (1904). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *L'année psychologique*, *11*(1), 191-244. <https://doi.org/10.3406/psy.1904.3675>
- Bishop, D. V. M., North, T., & Donlan, C. (1995). GENETIC BASIS OF SPECIFIC LANGUAGE IMPAIRMENT : EVIDENCE FROM A TWIN STUDY. *Developmental Medicine & Child Neurology*, *37*(1), 56-71. <https://doi.org/10.1111/j.1469-8749.1995.tb11932.x>
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition : A Longitudinal Study and an Intervention. *Child Development*, *78*(1), 246-263. <https://doi.org/10.1111/j.1467-8624.2007.00995.x>
- Blanchard, S., & Berger, S. (1994). Valeurs prédictives d'épreuves psychopédagogiques en français et en mathématiques utilisées en classe de 6e. In M. Huteau, *Les techniques psychologiques d'évaluation des personnes* (E.A.P., p. 595-598).
- Blanchard, S., Lieury, A., Le Cam, M., & Rocher, T. (2013). Motivation et sentiment d'efficacité personnelle chez 30 000 élèves de 6e du collège français. *Bulletin de psychologie*, Numéro 523(1), 23. <https://doi.org/10.3917/bupsy.523.0023>
- Blotière, P.-O., Miranda, S., Weill, A., Mikaeloff, Y., Peyre, H., Ramus, F., Mahmoud, Z., Coste, J., & Dray-Spira, R. (2020). Risk of early neurodevelopmental outcomes associated with prenatal exposure to the antiepileptic drugs most commonly used during pregnancy : A French nationwide population-based cohort study. *BMJ Open*, *10*(6), e034829. <https://doi.org/10.1136/bmjopen-2019-034829>
- Boets, B., & De Smedt, B. (2010). Single-digit arithmetic in children with dyslexia. *Dyslexia*, *16*(2), 183-191. <https://doi.org/10.1002/dys.403>
- Bolger, N., & Kellaghan, T. (1990). Method of Measurement and Gender Differences in Scholastic Achievement. *Journal of Educational Measurement*, *27*(2), 165-174. <https://doi.org/10.1111/j.1745-3984.1990.tb00740.x>
- Bollen, K. A. (1989). *Structural Equations with Latent Variables : Bollen/Structural Equations with Latent Variables*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118619179>
- Boonen, A. J. H., van der Schoot, M., van Wesel, F., de Vries, M. H., & Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade students. *Contemporary Educational Psychology*, *38*(3), 271-279. <https://doi.org/10.1016/j.cedpsych.2013.05.001>

- Borella, E., Carretti, B., & Pelegrina, S. (2010). The Specific Role of Inhibition in Reading Comprehension in Good and Poor Comprehenders. *Journal of Learning Disabilities, 43*(6), 541-552. <https://doi.org/10.1177/0022219410371676>
- Borghans, L., Golsteyn, B. H. H., Heckman, J. J., & Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences, 113*(47), 13354-13359. <https://doi.org/10.1073/pnas.1601135113>
- Bosse, M., Chaves, N., Largy, P., & Valdois, S. (2015). Orthographic learning during reading : The role of whole-word visual processing: ORTHOGRAPHIC LEARNING AND VISUAL PROCESSING. *Journal of Research in Reading, 38*(2), 141-158. <https://doi.org/10.1111/j.1467-9817.2012.01551.x>
- Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia : The visual attention span deficit hypothesis. *Cognition, 104*(2), 198-230. <https://doi.org/10.1016/j.cognition.2006.05.009>
- Bosse, M.-L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance : A cross-sectional study. *Journal of Research in Reading, 32*(2), 230-253. <https://doi.org/10.1111/j.1467-9817.2008.01387.x>
- Bourgeois, L. (2017, juin 7). *Enfant précoce : Est-ce vraiment une chance ?* Pratique.fr. <https://www.pratique.fr/actu/enfant-precoce-est-ce-vraiment-une-chance-9073079.html>
- Bradley, R. H., & Caldwell, B. M. (1984). The HOME Inventory and family demographics. *Developmental Psychology, 20*(2), 315-320. <https://doi.org/10.1037/0012-1649.20.2.315>
- Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Garcia Coll, C. (2001). The Home Environments of Children in the United States Part II: Relations with Behavioral Development through Age Thirteen. *Child Development, 72*(6), 1868-1886. <https://doi.org/10.1111/1467-8624.t01-1-00383>
- Bray, M. (2011). *The challenge of shadow education : Private tutoring and its implications for policy makers in the european union*. Technical report, European Commission.
- Breda, T., & Ly, S. T. (2015). Professors in Core Science Fields Are Not Always Biased against Women : Evidence from France. *American Economic Journal: Applied Economics, 7*(4), 53-75. <https://doi.org/10.1257/app.20140022>
- Breslau, J., Miller, E., Breslau, N., Bohnert, K., Lucia, V., & Schweitzer, J. (2009a). The Impact of Early Behavior Disturbances on Academic Achievement in High School. *PEDIATRICS, 123*(6), 1472-1476. <https://doi.org/10.1542/peds.2008-1406>
- Breslau, J., Miller, E., Breslau, N., Bohnert, K., Lucia, V., & Schweitzer, J. (2009b). The Impact of Early Behavior Disturbances on Academic Achievement in High School. *PEDIATRICS, 123*(6), 1472-1476. <https://doi.org/10.1542/peds.2008-1406>

- Britton, J. R., Britton, H. L., & Gronwaldt, V. (2006). Breastfeeding, sensitivity, and attachment. *Pediatrics*, *118*(5), e1436-1443. <https://doi.org/10.1542/peds.2005-2916>
- Brosseau-Liard, P. E., & Savalei, V. (2014). Adjusting Incremental Fit Indices for Nonnormality. *Multivariate Behavioral Research*, *49*(5), 460-470. <https://doi.org/10.1080/00273171.2014.933697>
- Brosseau-Liard, P. E., Savalei, V., & Li, L. (2012). An Investigation of the Sample Performance of Two Nonnormality Corrections for RMSEA. *Multivariate Behavioral Research*, *47*(6), 904-930. <https://doi.org/10.1080/00273171.2012.715252>
- Bub, K. L., McCartney, K., & Willett, J. B. (2007). Behavior problem trajectories and first-grade cognitive ability and achievement skills : A latent growth curve analysis. *Journal of Educational Psychology*, *99*(3), 653-670. <https://doi.org/10.1037/0022-0663.99.3.653>
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-Term Memory, Working Memory, and Executive Functioning in Preschoolers : Longitudinal Predictors of Mathematical Achievement at Age 7 Years. *Developmental Neuropsychology*, *33*(3), 205-228. <https://doi.org/10.1080/87565640801982312>
- Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability : Inhibition, Switching, and Working Memory. *Developmental Neuropsychology*, *19*(3), 273-293. https://doi.org/10.1207/S15326942DN1903_3
- Burden, M. J., Jacobson, S. W., Sokol, R. J., & Jacobson, J. L. (2005). Effects of Prenatal Alcohol Exposure on Attention and Working Memory at 7.5 Years of Age: *Alcoholism: Clinical & Experimental Research*, *29*(3), 443-452. <https://doi.org/10.1097/01.ALC.0000156125.50577.EC>
- Burke, M. A., & Sass, T. R. (2013). Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, *31*(1), 51-82. <https://doi.org/10.1086/666653>
- Butterfuss, R., & Kendeou, P. (2018). The Role of Executive Functions in Reading Comprehension. *Educational Psychology Review*, *30*(3), 801-826. <https://doi.org/10.1007/s10648-017-9422-6>
- Byrd, A. L., & Manuck, S. B. (2014). MAOA, Childhood Maltreatment, and Antisocial Behavior : Meta-analysis of a Gene-Environment Interaction. *Biological Psychiatry*, *75*(1), 9-17. <https://doi.org/10.1016/j.biopsych.2013.05.004>
- Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., Wadsworth, S., & DeFries, J. C. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood : Continuity and change from preschool to Grade 2. *Journal of Neurolinguistics*, *22*(3), 219-236. <https://doi.org/10.1016/j.jneuroling.2008.09.003>

- Cai, X., Lu, Y., Pan, J., & Zhong, S. (2018). Gender Gap under Pressure : Evidence from China's National College Entrance Examination. *The Review of Economics and Statistics*, *101*(2), 249–263. https://doi.org/10.1162/rest_a_00749
- Calero, M. D., García-Martín, M. B., Jiménez, M. I., Kazén, M., & Araque, A. (2007). Self-regulation advantage for high-IQ children : Findings from a research study. *Learning and Individual Differences*, *17*(4), 328-343. <https://doi.org/10.1016/j.lindif.2007.03.012>
- Calvin, C. M., Deary, I. J., Webbink, D., Smith, P., Fernandes, C., Lee, S. H., Luciano, M., & Visscher, P. M. (2012). Multivariate Genetic Analyses of Cognition and Academic Achievement from Two Population Samples of 174,000 and 166,000 School Children. *Behavior Genetics*, *42*(5), 699-710. <https://doi.org/10.1007/s10519-012-9549-7>
- Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., & Morrison, F. J. (2012). Fine motor skills and executive function both contribute to kindergarten achievement. *Child Development*, *83*(4), 1229-1244. <https://doi.org/10.1111/j.1467-8624.2012.01768.x>
- Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia*, *39*(11), 1224-1239. [https://doi.org/10.1016/S0028-3932\(01\)00035-5](https://doi.org/10.1016/S0028-3932(01)00035-5)
- Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The Foundations of Spelling Ability : Evidence from a 3-Year Longitudinal Study. *Journal of Memory and Language*, *45*(4), 751-774. <https://doi.org/10.1006/jmla.2000.2785>
- Carlson, A. G., Rowe, E., & Curby, T. W. (2013). Disentangling Fine Motor Skills' Relations to Academic Achievement : The Relative Contributions of Visual-Spatial Integration and Visual-Motor Coordination. *The Journal of Genetic Psychology*, *174*(5), 514-533. <https://doi.org/10.1080/00221325.2012.717122>
- Carman, C. A. (2013). Comparing Apples and Oranges : Fifteen Years of Definitions of Giftedness in Research. *Journal of Advanced Academics*, *24*(1), 52-70. <https://doi.org/10.1177/1932202X12472602>
- Carroll, J. M., Maughan, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders : Evidence for comorbidity. *Journal of Child Psychology and Psychiatry*, *46*(5), 524-532. <https://doi.org/10.1111/j.1469-7610.2004.00366.x>
- Casalis, S., & Louis-Alexandre, M.-F. (2000). Morphological analysis, phonological analysis and learning to read French : A longitudinal study. *Reading and Writing*, *12*(3/4), 303-335. <https://doi.org/10.1023/A:1008177205648>
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A., & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science (New York, N.Y.)*, *297*(5582), 851-854. <https://doi.org/10.1126/science.1072290>

- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars : Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence : A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22. <https://doi.org/10.1037/h0046743>
- Chamorro-Premuzic, T., Harlaar, N., Greven, C. U., & Plomin, R. (2010). More than just IQ : A longitudinal examination of self-perceived abilities as predictors of academic performance in a large sample of UK twins. *Intelligence*, 38(4), 385-392. <https://doi.org/10.1016/j.intell.2010.05.002>
- Chartier, P. (2012). *Evaluer les capacités de raisonnement avec les tests RCC - Raisonnement sur cartes de Chartier*. Eurotests Editions.
- Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of Adopted and Nonadopted Individuals Reveals Gene–Environment Interplay for Education in the UK Biobank. *Psychological Science*, 31(5), 582-591. <https://doi.org/10.1177/0956797620904450>
- Cheung, C. S.-S., & Pomerantz, E. M. (2012). Why does parents' involvement enhance children's achievement? The role of parent-oriented motivation. *Journal of Educational Psychology*, 104(3), 820-832. <https://doi.org/10.1037/a0027183>
- Chiappe, P., Siegel, L. S., & Hasher, L. (2000). Working memory, inhibitory control, and reading disability. *Memory & Cognition*, 28(1), 8-17. <https://doi.org/10.3758/BF03211570>
- Chinello, A., Cattani, V., Bonfiglioli, C., Dehaene, S., & Piazza, M. (2013). Objects, numbers, fingers, space : Clustering of ventral and dorsal functions in young children and adults. *Developmental Science*, 16(3), 377-393. <https://doi.org/10.1111/desc.12028>
- Chiodo, L. M., Da Costa, D. E., Hannigan, J. H., Covington, C. Y., Sokol, R. J., Janisse, J., Greenwald, M., Ager, J., & Delaney-Black, V. (2010). The Impact of Maternal Age on the Effects of Prenatal Alcohol Exposure on Attention : ALCOHOL AND ATTENTION MODERATED BY MATERNAL AGE. *Alcoholism: Clinical and Experimental Research*, 34(10), 1813-1821. <https://doi.org/10.1111/j.1530-0277.2010.01269.x>
- Christensen, J., Grønberg, T. K., Sørensen, M. J., Schendel, D., Parner, E. T., Pedersen, L. H., & Vestergaard, M. (2013). Prenatal Valproate Exposure and Risk of Autism Spectrum Disorders and Childhood Autism. *JAMA*, 309(16), 1696. <https://doi.org/10.1001/jama.2013.2270>
- Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement : Evidence from the ECLS-K. *Economics of Education Review*, 28(4), 415-427. <https://doi.org/10.1016/j.econedurev.2008.09.003>
- Clifford, A., Lang, L., & Chen, R. (2012). Effects of maternal cigarette smoking during pregnancy on cognitive parameters of children and young adults : A literature

- review. *Neurotoxicology and Teratology*, 34(6), 560-570.
<https://doi.org/10.1016/j.ntt.2012.09.004>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (Routledge).
- Cohen, L., & Dehaene, S. (2000). CALCULATING WITHOUT READING : UNSUSPECTED RESIDUAL ABILITIES IN PURE ALEXIA. *Cognitive Neuropsychology*, 17(6), 563-583. <https://doi.org/10.1080/02643290050110656>
- Colonat, A. (2018, janvier 18). *Comment peut-on être surdoué et pourtant en échec scolaire? - Science & Vie*. <https://www.science-et-vie.com/questions-reponses/comment-peut-on-etre-surdoue-et-pourtant-en-echec-scolaire-10327>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <https://doi.org/10.1037/0033-295x.108.1.204>
- Costa, A. J., Silva, J. B. L., Pinheiro-Chagas, P., Krinzinger, H., Lohnemann, J., Willmes, K., Wood, G., & Haase, V. G. (2011). A hand full of numbers : A role for offloading in arithmetics learning? *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00368>
- Costello, E. J., Worthman, C., Erkanli, A., & Angold, A. (2007). Prediction From Low Birth Weight to Female Adolescent Depression : A Test of Competing Hypotheses. *Archives of General Psychiatry*, 64(3), 338. <https://doi.org/10.1001/archpsyc.64.3.338>
- Côte, S. (2005, janvier 29). *L'enfant prometteur et l'école*. Association Française pour les Enfants Précoces.
- Courchia, B., Berkovits, M. D., & Bauer, C. R. (2019). Cognitive impairment among extremely low birthweight preterm infants from 1980 to present day. *Journal of Perinatology*, 39(8), 1098-1104. <https://doi.org/10.1038/s41372-019-0414-x>
- Couto, J. M., Gomez, L., Wigg, K., Ickowicz, A., Pathare, T., Malone, M., Kennedy, J. L., Schachar, R., & Barr, C. L. (2009). Association of Attention-Deficit/Hyperactivity Disorder with a Candidate Region for Reading Disabilities on Chromosome 6p. *Biological Psychiatry*, 66(4), 368-375. <https://doi.org/10.1016/j.biopsych.2009.02.016>
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2), 31-47. <https://doi.org/10.1257/aer.97.2.31>
- de Kieviet, J. F., Piek, J. P., Aarnoudse-Moens, C. S., & Oosterlaan, J. (2009). Motor Development in Very Preterm and Very Low-Birth-Weight Children From Birth to Adolescence : A Meta-analysis. *JAMA*, 302(20), 2235. <https://doi.org/10.1001/jama.2009.1708>
- De Smedt, B., & Boets, B. (2010). Phonological processing and arithmetic fact retrieval : Evidence from developmental dyslexia. *Neuropsychologia*, 48(14), 3973-3981. <https://doi.org/10.1016/j.neuropsychologia.2010.10.018>

- De Smedt, B., Taylor, J., Archibald, L., & Ansari, D. (2010). How is phonological processing related to individual differences in children's arithmetic skills? : Phonological processing and arithmetic. *Developmental Science*, *13*(3), 508-520. <https://doi.org/10.1111/j.1467-7687.2009.00897.x>
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21. <https://doi.org/10.1016/j.intell.2006.02.001>
- Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*(83-120).
- Dehaene, S., & Cohen, L. (1997). Cerebral Pathways for Calculation : Double Dissociation between Rote Verbal and Quantitative Knowledge of Arithmetic. *Cortex*, *33*(2), 219-250. [https://doi.org/10.1016/S0010-9452\(08\)70002-9](https://doi.org/10.1016/S0010-9452(08)70002-9)
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). THREE PARIETAL CIRCUITS FOR NUMBER PROCESSING. *Cognitive Neuropsychology*, *20*(3-6), 487-506. <https://doi.org/10.1080/02643290244000239>
- DeMars, C. E., Bashkov, B. M., & Socha, A. B. (2013). The role of gender in test-taking motivation under low-stakes conditions. *Research and Practice in Assessment*, *8*, 69–82.
- Demir, Ö. E., Prado, J., & Booth, J. R. (2015). Parental socioeconomic status and the neural basis of arithmetic : Differential relations to verbal and visuo-spatial representations. *Developmental Science*, *18*(5), 799-814. <https://doi.org/10.1111/desc.12268>
- Der, G., Batty, G. D., & Deary, I. J. (2009). The association between IQ in adolescence and a range of health outcomes at 40 in the 1979 US National Longitudinal Study of Youth. *Intelligence*, *37*(6), 573-580. <https://doi.org/10.1016/j.intell.2008.12.002>
- DiStefano, C., Zhu, M., & Míndrilă, D. (2019). Understanding and Using Factor Scores : Considerations for the Applied Researcher. *Practical Assessment, Research, and Evaluation*, *14*(1). <https://doi.org/10.7275/da8t-4g52>
- Dobbs, J., Doctoroff, G. L., Fisher, P. H., & Arnold, D. H. (2006). The association between preschool children's socio-emotional functioning and their mathematical skills. *Journal of Applied Developmental Psychology*, *27*(2), 97-108. <https://doi.org/10.1016/j.appdev.2005.12.008>
- Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences*, *108*(19), 7716-7720. <https://doi.org/10.1073/pnas.1018601108>
- Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind : The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*(2), 439-451. <https://doi.org/10.1037/a0026280>

- Duckworth, A. L., & Seligman, M. E. P. (2005). Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. *Psychological Science, 16*(12), 939-944. <https://doi.org/10.1111/j.1467-9280.2005.01641.x>
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge : Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*(1), 198-208. <https://doi.org/10.1037/0022-0663.98.1.198>
- Duckworth, A. L., Shulman, E. P., Mastronarde, A. J., Patrick, S. D., Zhang, J., & Druckman, J. (2015). Will not want : Self-control rather than motivation explains the female advantage in report card grades. *Learning and Individual Differences, 39*, 13-23. <https://doi.org/10.1016/j.lindif.2015.02.006>
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic Deprivation and Early Childhood Development. *Child Development, 65*(2), 296-318. <https://doi.org/10.1111/j.1467-8624.1994.tb00752.x>
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91*(2), 113-136. <https://doi.org/10.1016/j.jecp.2005.01.003>
- Durand, V. N., Loe, I. M., Yeatman, J. D., & Feldman, H. M. (2013). Effects of early language, speech, and cognition on later reading : A mediation analysis. *Frontiers in Psychology, 4*. <https://doi.org/10.3389/fpsyg.2013.00586>
- Ehrenstein, V. (2009). Association of Apgar scores with death and neurologic disability. *Clinical Epidemiology, 45*. <https://doi.org/10.2147/CLEP.S4782>
- Ehri, L. C. (2000). Learning To Read and Learning To Spell : Two Sides of a Coin. *Topics in Language Disorders, 20*(3), 19-36. <https://doi.org/10.1097/00011363-200020030-00005>
- Eklöf, H. (2007). Test-Taking Motivation and Mathematics Performance in TIMSS 2003. *International Journal of Testing, 7*(3), 311-326. <https://doi.org/10.1080/15305050701438074>
- Elbro, C., & Scarborough, H. (2004). Early Identification. *Handbook of Children's Literacy, 339-359*. https://doi.org/10.1007/978-94-017-1731-1_19
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics : A meta-analysis. *Psychological Bulletin, 136*(1), 103-127. <https://doi.org/10.1037/a0018053>
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The Contribution of Executive Functions to Emergent Mathematic Skills

- in Preschool Children. *Developmental Neuropsychology*, 26(1), 465-486.
https://doi.org/10.1207/s15326942dn2601_6
- Facoetti, A., Corradi, N., Ruffino, M., Gori, S., & Zorzi, M. (2010). Visual spatial attention and speech segmentation are both impaired in preschoolers at familial risk for developmental dyslexia. *Dyslexia*, 16(3), 226-239.
<https://doi.org/10.1002/dys.413>
- Falch, T., & Naper, L. R. (2013). Educational evaluation schemes and gender gaps in student achievement. *Economics of Education Review*, 36, 12-25.
<https://doi.org/10.1016/j.econedurev.2013.05.002>
- Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist*, 43(2), 95-103. <https://doi.org/10.1037/0003-066X.43.2.95>
- Feldhusen, J. F., & Kroll, M. D. (1991). Boredom or Challenge for the Academically Talented in School. *Gifted Education International*, 7(2), 80-81.
<https://doi.org/10.1177/026142949100700207>
- Fergusson, D. M., & Horwood, L. J. (1995). Early disruptive behavior, IQ, and later school achievement and delinquent behavior. *Journal of Abnormal Child Psychology*, 23(2), 183-199. <https://doi.org/10.1007/BF01447088>
- Flak, A. L., Su, S., Bertrand, J., Denny, C. H., Kesmodel, U. S., & Cogswell, M. E. (2014). The Association of Mild, Moderate, and Binge Prenatal Alcohol Exposure and Child Neuropsychological Outcomes : A Meta-Analysis. *Alcoholism: Clinical and Experimental Research*, 38(1), 214-226. <https://doi.org/10.1111/acer.12214>
- Florit, E., & Cain, K. (2011). The Simple View of Reading : Is It Valid for Different Types of Alphabetic Orthographies? *Educational Psychology Review*, 23(4), 553-576. <https://doi.org/10.1007/s10648-011-9175-6>
- Fluss, J., Ziegler, J. C., Warszawski, J., Ducot, B., Richard, G., & Billard, C. (2009). Poor Reading in French Elementary School : The Interplay of Cognitive, Behavioral, and Socioeconomic Factors: *Journal of Developmental & Behavioral Pediatrics*, 30(3), 206-216. <https://doi.org/10.1097/DBP.0b013e3181a7ed6c>
- Flynn, J. R. (1991). *Asian Americans : Achievement beyond IQ*. L. Erlbaum Associates.
- Frankenburg, W. K., & Coons, C. E. (1986). Home screening questionnaire : Its validity in assessing home environment. *The Journal of Pediatrics*, 108(4), 624-626.
[https://doi.org/10.1016/S0022-3476\(86\)80853-8](https://doi.org/10.1016/S0022-3476(86)80853-8)
- Fried, P. A., Watkinson, B., & Siegel, L. S. (1997). Reading and language in 9- to 12-year olds prenatally exposed to cigarettes and marijuana. *Neurotoxicology and Teratology*, 19(3), 171-183. [https://doi.org/10.1016/S0892-0362\(97\)00015-9](https://doi.org/10.1016/S0892-0362(97)00015-9)
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The Contributions of Numerosity and Domain-General Abilities to School Readiness : Numerosity and Domain-General Abilities. *Child Development*, 81(5), 1520-1533. <https://doi.org/10.1111/j.1467-8624.2010.01489.x>

- Gale, C. R., Hatch, S. L., Batty, G. D., & Deary, I. J. (2009). Intelligence in childhood and risk of psychological distress in adulthood : The 1958 National Child Development Survey and the 1970 British Cohort Study. *Intelligence*, *37*(6), 592-599. <https://doi.org/10.1016/j.intell.2008.09.002>
- Gallagher, A., Levin, J., & Cahalan, C. (2002). Cognitive patterns of gender differences on mathematics admissions tests. *ETS Research Report Series*, *2002*(2), i-30. <https://doi.org/10.1002/j.2333-8504.2002.tb01886.x>
- Gashaj, V., Oberer, N., Mast, F. W., & Roebers, C. M. (2019). The Relation Between Executive Functions, Fine Motor Skills, and Basic Numerical Skills and Their Relevance for Later Mathematics Achievement. *Early Education and Development*, *30*(7), 913-926. <https://doi.org/10.1080/10409289.2018.1539556>
- Geary, D. C. (2000). From infancy to adulthood : The development of numerical abilities. *European Child & Adolescent Psychiatry*, *9*(S2), S11-S16. <https://doi.org/10.1007/s007870070004>
- Geary, D. C. (2013). Early Foundations for Mathematics Learning and Their Relations to Learning Disabilities. *Current Directions in Psychological Science*, *22*(1), 23-27. <https://doi.org/10.1177/0963721412469398>
- Gentaz, E., Sprenger-Charolles, L., & Theurel, A. (2015). Differences in the Predictors of Reading Comprehension in First Graders from Low Socio-Economic Status Families with Either Good or Poor Decoding Skills. *PLOS ONE*, *10*(3), e0119581. <https://doi.org/10.1371/journal.pone.0119581>
- Germanò, E., Gagliano, A., & Curatolo, P. (2010). Comorbidity of ADHD and Dyslexia. *Developmental Neuropsychology*, *35*(5), 475-493. <https://doi.org/10.1080/87565641.2010.494748>
- Gignac, G. E. (2018). A moderate financial incentive can increase effort, but not intelligence test performance in adult volunteers. *British Journal of Psychology*, *109*(3), 500-516. <https://doi.org/10.1111/bjop.12288>
- Gilman, S. E., Gardener, H., & Buka, S. L. (2008). Maternal Smoking during Pregnancy and Children's Cognitive and Physical Development : A Causal Risk Factor? *American Journal of Epidemiology*, *168*(5), 522-531. <https://doi.org/10.1093/aje/kwn175>
- Girard, L.-C., Doyle, O., & Tremblay, R. E. (2018). Breastfeeding and externalising problems : A quasi-experimental design with a national cohort. *European Child & Adolescent Psychiatry*, *27*(7), 877-884. <https://doi.org/10.1007/s00787-017-1085-9>
- Goigoux, R., Cèbe, S., & Pironom, J. (2016). Les facteurs explicatifs des performances en lecture-compréhension à la fin du cours préparatoire. *Revue française de pédagogie*, *196*, 67-84. <https://doi.org/10.4000/rfp.5076>
- Goisis, A., Schneider, D. C., & Myrskylä, M. (2017). The reversing association between advanced maternal age and child cognitive ability : Evidence from three UK birth

- cohorts. *International Journal of Epidemiology*, 46(3), 850-859.
<https://doi.org/10.1093/ije/dyw354>
- Goldschmidt, L., Richardson, G. A., Cornelius, M. D., & Day, N. L. (2004). Prenatal marijuana and alcohol exposure and academic achievement at age 10. *Neurotoxicology and Teratology*, 26(4), 521-532.
<https://doi.org/10.1016/j.ntt.2004.04.003>
- Gomez, A., Piazza, M., Jobert, A., Dehaene-Lambertz, G., Dehaene, S., & Huron, C. (2015). Mathematical difficulties in developmental coordination disorder : Symbolic and nonsymbolic number processing. *Research in Developmental Disabilities*, 43-44, 167-178. <https://doi.org/10.1016/j.ridd.2015.06.011>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire : A Research Note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586.
<https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337-1345. <https://doi.org/10.1097/00004583-200111000-00015>
- Gordon, M. T. (1976). A Different View of the IQ-Achievement Gap. *Sociology of Education*, 49(1), 4. <https://doi.org/10.2307/2112387>
- Gori, S., Seitz, A. R., Ronconi, L., Franceschini, S., & Facoetti, A. (2016). Multiple Causal Links Between Magnocellular–Dorsal Pathway Deficit and Developmental Dyslexia. *Cerebral Cortex*, 26(11), 4356-4369.
<https://doi.org/10.1093/cercor/bhv206>
- Gough, P. B., Hoover, V. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In *Reading comprehension difficulties : Processes and intervention* (Lawrence Erlbaum Associates Publishers, p. 1-13).
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7(1), 6-10.
<https://doi.org/10.1177/074193258600700104>
- Greenberg, M. T., Lengua, L. J., Coie, J. D., & Pinderhughes, E. E. (1999). Predicting developmental outcomes at school entry using a multiple-risk model : Four American communities. The Conduct Problems Prevention Research Group. *Developmental Psychology*, 35(2), 403-417.
- Greven, C. U., Harlaar, N., Kovas, Y., Chamorro-Premuzic, T., & Plomin, R. (2009). More Than Just IQ : School Achievement Is Predicted by Self-Perceived Abilities—But for Genetic Rather Than Environmental Reasons. *Psychological Science*, 20(6), 753-762. <https://doi.org/10.1111/j.1467-9280.2009.02366.x>
- Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010a). Fine motor skills and early comprehension of the world : Two new school readiness indicators. *Developmental Psychology*, 46(5), 1008-1017.
<https://doi.org/10.1037/a0020104>

- Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010b). Fine motor skills and early comprehension of the world : Two new school readiness indicators. *Developmental Psychology*, *46*(5), 1008-1017. <https://doi.org/10.1037/a0020104>
- Grobman, J. (2006). Underachievement in Exceptionally Gifted Adolescents and Young Adults : A Psychiatrist's View. *Journal of Secondary Gifted Education*, *17*(4), 199-210. <https://doi.org/10.4219/jsge-2006-408>
- Guénolé, F., Louis, J., Creveuil, C., Baleyte, J.-M., Montlahuc, C., Fournieret, P., & Revol, O. (2013). Behavioral Profiles of Clinically Referred Children with Intellectual Giftedness. *BioMed Research International*, *2013*, 1-7. <https://doi.org/10.1155/2013/540153>
- Guez, A., Panaiotis, T., Peyre, H., & Ramus, F. (2018). Predictors of the IQ-achievement gap in France : A longitudinal analysis. *Intelligence*, *69*, 104-116. <https://doi.org/10.1016/j.intell.2018.05.008>
- Guhn, M., Emerson, S. D., Mahdavian, D., & Gadermann, A. M. (2020). Associations of Birth Factors and Socio-Economic Status with Indicators of Early Emotional Development and Mental Health in Childhood : A Population-Based Linkage Study. *Child Psychiatry & Human Development*, *51*(1), 80-93. <https://doi.org/10.1007/s10578-019-00912-6>
- Guyon, N., & Huillery, E. (2016). *Biased Aspirations and Social Inequality at School : Evidence from French Teenagers*. LIEPP Working Paper n°44.
- Hahn, R. A., & Truman, B. I. (2015). Education Improves Public Health and Promotes Health Equity. *International Journal of Health Services*, *45*(4), 657-678. <https://doi.org/10.1177/0020731415585986>
- Halonen, A., Aunola, K., Ahonen, T., & Nurmi, J.-E. (2006). The role of learning to read in the development of problem behaviour : A cross-lagged longitudinal study. *British Journal of Educational Psychology*, *76*(3), 517-534. <https://doi.org/10.1348/000709905X51590>
- Halpern, D. F. (2002). Sex Differences in Achievement Scores : Can We Design Assessments that are Fair, Meaningful, and Valid for Girls and Boys? *Issues in Education*, *8*(1), 2.
- Hanna, R. N., & Linden, L. L. (2012). Discrimination in Grading. *American Economic Journal: Economic Policy*, *4*(4), 146-168. <https://doi.org/10.1257/pol.4.4.146>
- Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, *18*(5), 527-544. <https://doi.org/10.1002/jae.741>
- Harden, K. P., Domingue, B. W., Belsky, D. W., Boardman, J. D., Crosnoe, R., Malanchini, M., Nivard, M., Tucker-Drob, E. M., & Harris, K. M. (2020). Genetic associations with mathematics tracking and persistence in secondary school. *Npj Science of Learning*, *5*(1). <https://doi.org/10.1038/s41539-020-0060-2>

- Hart, S. A., Petrill, S. A., Willcutt, E., Thompson, L. A., Schatschneider, C., Deater-Deckard, K., & Cutting, L. E. (2010). Exploring How Symptoms of Attention-Deficit/Hyperactivity Disorder Are Related to Reading and Mathematics Performance : General Genes, General Environments. *Psychological Science*, *21*(11), 1708-1715. <https://doi.org/10.1177/0956797610386617>
- Hausmann, R., Tyson, L. D., Zahidi, S., & World Economic Forum. (2010). *The global gender gap report 2010*. http://www3.weforum.org/docs/WEF_GenderGap_Report_2011.pdf
- Havron, N., Ramus, F., Heude, B., Forhan, A., Cristia, A., Peyre, H., the EDEN Mother-Child Cohort Study Group, Annesi-Maesano, I., Bernard, J. Y., Botton, J., Charles, M. A., Dargent-Molina, P., de Lauzon-Guillain, B., Ducimetière, P., De Agostini, M., Foliguet, B., Forhan, A., Fritel, X., Germa, A., ... Thiébauges, O. (2019). The Effect of Older Siblings on Language Development as a Function of Age Difference and Sex. *Psychological Science*, *30*(9), 1333-1343. <https://doi.org/10.1177/0956797619861436>
- Hawes, Z., Moss, J., Caswell, B., Seo, J., & Ansari, D. (2019). Relations between numerical, spatial, and executive function skills and mathematics achievement : A latent-variable approach. *Cognitive Psychology*, *109*, 68-90. <https://doi.org/10.1016/j.cogpsych.2018.12.002>
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science (New York, N.Y.)*, *269*(5220), 41-45.
- Heikkilä, K., Kelly, Y., Renfrew, M. J., Sacker, A., & Quigley, M. A. (2014). Breastfeeding and educational achievement at age 5 : Breastfeeding and early educational achievement. *Maternal & Child Nutrition*, *10*(1), 92-101. <https://doi.org/10.1111/j.1740-8709.2012.00402.x>
- Heude, B., Forhan, A., Slama, R., Douhaud, L., Bedel, S., Saurel-Cubizolles, M.-J., Hankard, R., Thiébauges, O., De Agostini, M., Annesi-Maesano, I., Kaminski, M., & Charles, M.-A. (2016). Cohort Profile : The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *International Journal of Epidemiology*, *45*(2), 353-363. <https://doi.org/10.1093/ije/dyv151>
- Hill, N. E., & Craft, S. A. (2003). Parent-school involvement and school performance : Mediated pathways among socioeconomically comparable African American and Euro-American families. *Journal of Educational Psychology*, *95*(1), 74-83. <https://doi.org/10.1037/0022-0663.95.1.74>
- Hjetland, H. N. (2018). *Pathways to reading comprehension : A longitudinal study from 4 to 9 years of age*. Journal of Educational Psychology; US: American Psychological Association. <https://doi.org/10.1037/edu0000321>
- Hjetland, H. N., Brinchmann, E. I., Scherer, R., Hulme, C., & Melby-Lervåg, M. (2020). Preschool pathways to reading comprehension : A systematic meta-analytic review. *Educational Research Review*, 100323. <https://doi.org/10.1016/j.edurev.2020.100323>

- Holmes, J., & Adams, J. W. (2006). Working Memory and Children's Mathematical Skills : Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339-366. <https://doi.org/10.1080/01443410500341056>
- Hood, M., Conlon, E., & Andrews, G. (2008). Preschool home literacy practices and children's literacy development : A longitudinal analysis. *Journal of Educational Psychology, 100*(2), 252-271. <https://doi.org/10.1037/0022-0663.100.2.252>
- Hoover, V. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2*(2). <https://doi.org/10.1007/BF00401799>
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*(5), 253-270. <https://doi.org/10.1037/h0023816>
- Horta, B. L., Loret de Mola, C., & Victora, C. G. (2015). Breastfeeding and intelligence : A systematic review and meta-analysis. *Acta Paediatrica, 104*, 14-19. <https://doi.org/10.1111/apa.13139>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis : Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The Causal Role of Phoneme Awareness and Letter-Sound Knowledge in Learning to Read : Combining Intervention Studies With Mediation Analyses. *Psychological Science, 23*(6), 572-577. <https://doi.org/10.1177/0956797611435921>
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance : A meta-analysis. *Psychological Bulletin, 107*(2), 139-155. <https://doi.org/10.1037/0033-2909.107.2.139>
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability : A meta-analysis. *Psychological Bulletin, 104*(1), 53-69. <https://doi.org/10.1037/0033-2909.104.1.53>
- Ip, S., Chung, M., Raman, G., Chew, P., Magula, N., DeVine, D., Trikalinos, T., & Lau, J. (2007). Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence Report/Technology Assessment, 153*, 1-186.
- Jackson, P. S., & Peterson, J. (2003). Depressive Disorder in Highly Gifted Adolescents. *Journal of Secondary Gifted Education, 14*(3), 175-186. <https://doi.org/10.4219/jsge-2003-429>
- Jacobson, J. L., & Jacobson, S. W. (2002). Effects of Prenatal Alcohol Exposure on Child Development. *Alcohol Research & Health, 26*(4), 282-286.
- Johnson, W., McGue, M., & Iacono, W. G. (2007). Socioeconomic status and school grades : Placing their association in broader context in a sample of biological and adoptive families. *Intelligence, 35*(6), 526-541. <https://doi.org/10.1016/j.intell.2006.09.006>

- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters : Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850-867. <https://doi.org/10.1037/a0014939>
- Jurajda, Š., & Münich, D. (2011). Gender Gap in Performance under Competitive Pressure : Admissions to Czech Universities. *American Economic Review*, 101(3), 514-518. <https://doi.org/10.1257/aer.101.3.514>
- Jurecska, D. E., Chang, K. B. T., Peterson, M. A., Lee-Zorn, C. E., Merrick, J., & Sequeira, E. (2012). The poverty puzzle : The surprising difference between wealthy and poor students for self-efficacy and academic achievement. *International Journal of Adolescent Medicine and Health*, 24(4). <https://doi.org/10.1515/ijamh.2012.052>
- Karpinski, R. I., Kinase Kolb, A. M., Tetreault, N. A., & Borowski, T. B. (2018). High intelligence : A risk factor for psychological and physiological overexcitabilities. *Intelligence*, 66, 8-23. <https://doi.org/10.1016/j.intell.2017.09.001>
- Kemp, S. L., Kirk, U., & Korkman, M. (2001). *Essentials of NEPSY Assessment* (1st ed). Wiley.
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school : The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, 101(4), 765-778. <https://doi.org/10.1037/a0015956>
- Kersey, A. J., Braham, E. J., Csumitta, K. D., Libertus, M. E., & Cantlon, J. F. (2018). No intrinsic gender differences in children's earliest numerical abilities. *Npj Science of Learning*, 3(1). <https://doi.org/10.1038/s41539-018-0028-7>
- Kling, K. C., Nofle, E. E., & Robins, R. W. (2013). Why Do Standardized Tests Underpredict Women's Academic Performance? The Role of Conscientiousness. *Social Psychological and Personality Science*, 4(5), 600-606. <https://doi.org/10.1177/1948550612469038>
- Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation : Evidence from the operational momentum effect. *Attention, Perception, & Psychophysics*, 71(4), 803-821. <https://doi.org/10.3758/APP.71.4.803>
- Korkman, M., Kirk, U., & Kemp, S. L. (2003). *Nepsy Bilan Neuropsychologique de l'enfant*. ECPA (Editions du Centre de Psychologie Appliquée).
- Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and Numeracy Are More Heritable Than Intelligence in Primary School. *Psychological Science*, 24(10), 2048-2056. <https://doi.org/10.1177/0956797613486982>
- Kramer, M. S. (2008). Breastfeeding and Child Cognitive Development : New Evidence From a Large Randomized Trial. *Archives of General Psychiatry*, 65(5), 578. <https://doi.org/10.1001/archpsyc.65.5.578>

- Krapohl, E., Hannigan, L. J., Pingault, J.-B., Patel, H., Kadeva, N., Curtis, C., Breen, G., Newhouse, S. J., Eley, T. C., O'Reilly, P. F., & Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences*, *114*(44), 11727-11732. <https://doi.org/10.1073/pnas.1707178114>
- Krapohl, Eva, & Plomin, R. (2016). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, *21*(3), 437-443. <https://doi.org/10.1038/mp.2015.2>
- Krapohl, Eva, Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B., Asbury, K., Harlaar, N., Kovas, Y., Dale, P. S., & Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences*, *111*(42), 15273-15278. <https://doi.org/10.1073/pnas.1408777111>
- Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation : A predictor of PISA's mathematical competence beyond intelligence and prior test achievement. *Learning and Individual Differences*, *43*, 140-148. <https://doi.org/10.1016/j.lindif.2015.08.026>
- Kuzujanakis, M. (2013, avril 10). The Misunderstood Face of Giftedness. *Huffington Post*. https://www.huffingtonpost.com/marianne-kuzujanakis/gifted-children_b_2948258.html
- Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten : Related spheres of influence? *Child Development*, *70*(6), 1373-1400. <https://doi.org/10.1111/1467-8624.00101>
- Lafontaine, D., & Monseur, C. (2009a). Les évaluations des performances en mathématiques sont-elles influencées par le sexe de l'élève? *Mesure et évaluation en éducation*, *32*(2), 71-98.
- Lafontaine, D., & Monseur, C. (2009b). Gender Gap in Comparative Studies of Reading Comprehension : To What Extent Do the Test Characteristics Make a Difference? *European Educational Research Journal*, *8*(1), 69-79. <https://doi.org/10.2304/eeerj.2009.8.1.69>
- Lancon, C., Martinelli, M., Michel, P., Debals, M., Auquier, P., Guedj, E., & Boyer, L. (2015). Comorbidités psychiatriques et qualité de vie chez les sujets adultes à haut potentiel intellectuel : Relations avec l'estime de soi. *La Presse Médicale*, *44*(5), e177-e184. <https://doi.org/10.1016/j.lpm.2014.11.018>
- Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography : An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150-161. <https://doi.org/10.1037/0022-0663.100.1.150>
- Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, *92*(10-11), 2083-2105. <https://doi.org/10.1016/j.jpubeco.2008.02.009>

- Lawson, G. M., Hook, C. J., & Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*(2), e12529. <https://doi.org/10.1111/desc.12529>
- Le Saint, R. (2017, mai 30). *Comment repérer les « surdoués » autour de vous au bureau*. BFMTV. <http://www.bfmtv.com/sante/comment-reperer-les-surdoues-autour-de-vous-au-bureau-1100508.html>
- Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., ... Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112-1121. <https://doi.org/10.1038/s41588-018-0147-3>
- Lee, K.-M. (2000). Cortical areas differentially involved in multiplication and subtraction : A functional magnetic resonance imaging study and correlation with a case of selective acalculia. *Annals of Neurology*, *48*(4), 657-661.
- Lee, K.-M., & Kang, S.-Y. (2002). Arithmetic operation and working memory : Differential suppression in dual tasks. *Cognition*, *83*(3), B63-B68. [https://doi.org/10.1016/S0010-0277\(02\)00010-0](https://doi.org/10.1016/S0010-0277(02)00010-0)
- LeFevre, J.-A., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S.-L., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in Grades 2 through 4. *Journal of Experimental Child Psychology*, *114*(2), 243-261. <https://doi.org/10.1016/j.jecp.2012.10.005>
- LeFevre, J.-A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition : Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 216-230. <https://doi.org/10.1037/0278-7393.22.1.216>
- Lehrl, S., Ebert, S., Blaurock, S., Rossbach, H.-G., & Weinert, S. (2020). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement*, *31*(1), 102-124. <https://doi.org/10.1080/09243453.2019.1618346>
- Leroy, N., Bressoux, P., Sarrazin, P., & Trouilloud, D. (2013). Un modèle sociocognitif des apprentissages scolaires : Style motivationnel de l'enseignant, soutien perçu des élèves et processus motivationnels. *Revue française de pédagogie*, *182*, 71-92. <https://doi.org/10.4000/rfp.4008>
- Lervåg, A., & Hulme, C. (2010). Predicting the Growth of Early Spelling Skills : Are There Heterogeneous Developmental Trajectories? *Scientific Studies of Reading*, *14*(6), 485-513. <https://doi.org/10.1080/10888431003623488>

- Letourneau, N. L., Duffett-Leger, L., Levac, L., Watson, B., & Young-Morris, C. (2013). Socioeconomic Status and Child Development : A Meta-Analysis. *Journal of Emotional and Behavioral Disorders, 21*(3), 211-224. <https://doi.org/10.1177/1063426611421007>
- Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic Status Modifies the Sex Difference in Spatial Skill. *Psychological Science, 16*(11), 841-845. <https://doi.org/10.1111/j.1467-9280.2005.01623.x>
- Li, M., Liu, D., Li, M., Dong, W., Huang, Y., & Chen, Q. (2018). Addition and Subtraction but Not Multiplication and Division Cause Shifts of Spatial Attention. *Frontiers in Human Neuroscience, 12*. <https://doi.org/10.3389/fnhum.2018.00183>
- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika, 73*(1), 13. <https://doi.org/10.2307/2336267>
- Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences, 25*, 126-133. <https://doi.org/10.1016/j.lindif.2013.02.001>
- Lieury, A. (1996). Mémoire encyclopédique et devenir scolaire. Étude longitudinale d'une cohorte sur les quatre années du collège français. *Psychologie et psychométrie, 17*(3), 33-44.
- Lieury, A. (2012). *Mémoire et réussite scolaire* (Dunod).
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance : A meta-analysis. *Psychological Bulletin, 136*(6), 1123-1135. <https://doi.org/10.1037/a0021276>
- Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(6), 466-474. <https://doi.org/10.1016/j.jaac.2017.03.013>
- Lubinski, D., & Benbow, C. P. (2006). Study of Mathematically Precocious Youth After 35 Years : Uncovering Antecedents for the Development of Math-Science Expertise. *Perspectives on Psychological Science, 1*(4), 316-345. <https://doi.org/10.1111/j.1745-6916.2006.00019.x>
- Luo, Z., Jose, P. E., Huntsinger, C. S., & Pigott, T. D. (2007a). Fine motor skills and mathematics achievement in East Asian American and European American kindergartners and first graders. *British Journal of Developmental Psychology, 25*(4), 595-614. <https://doi.org/10.1348/026151007X185329>
- Luo, Z., Jose, P. E., Huntsinger, C. S., & Pigott, T. D. (2007b). Fine motor skills and mathematics achievement in East Asian American and European American kindergartners and first graders. *British Journal of Developmental Psychology, 25*(4), 595-614. <https://doi.org/10.1348/026151007X185329>

- Lynn, R., & Mikk, J. (2009). SEX DIFFERENCES IN READING ACHIEVEMENT. *Trames. Journal of the Humanities and Social Sciences*, 13(1), 3. <https://doi.org/10.3176/tr.2009.1.01>
- Madigan, S., Browne, D., Racine, N., Mori, C., & Tough, S. (2019). Association Between Screen Time and Children’s Performance on a Developmental Screening Test. *JAMA Pediatrics*, 173(3), 244. <https://doi.org/10.1001/jamapediatrics.2018.5056>
- Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R., & Christakis, D. A. (2020). Associations Between Screen Use and Child Language Skills : A Systematic Review and Meta-analysis. *JAMA Pediatrics*. <https://doi.org/10.1001/jamapediatrics.2020.0327>
- Maeda, Y., & Yoon, S. Y. (2013). A Meta-Analysis on Gender Differences in Mental Rotation Ability Measured by the Purdue Spatial Visualization Tests : Visualization of Rotations (PSVT:R). *Educational Psychology Review*, 25(1), 69-94. <https://doi.org/10.1007/s10648-012-9215-x>
- Malanchini, M., Engelhardt, L. E., Grotzinger, A. D., Harden, K. P., & Tucker-Drob, E. M. (2019). “Same but different” : Associations between multiple aspects of self-regulation, cognition, and academic abilities. *Journal of Personality and Social Psychology*, 117(6), 1164-1188. <https://doi.org/10.1037/pspp0000224>
- Masseti, G. M., Lahey, B. B., Pelham, W. E., Loney, J., Ehrhardt, A., Lee, S. S., & Kipp, H. (2008). Academic Achievement Over 8 Years Among Children Who Met Modified Criteria for Attention-deficit/Hyperactivity Disorder at 4–6 Years of Age. *Journal of Abnormal Child Psychology*, 36(3), 399-410. <https://doi.org/10.1007/s10802-007-9186-4>
- Massonnié, J., Bianco, M., Lima, L., & Bressoux, P. (2019). Longitudinal predictors of reading comprehension in French at first grade : Unpacking the oral comprehension component of the simple view. *Learning and Instruction*, 60, 166-179. <https://doi.org/10.1016/j.learninstruc.2018.01.005>
- Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradović, J., Riley, J. R., Boelcke-Stennes, K., & Tellegen, A. (2005). Developmental Cascades : Linking Academic Achievement and Externalizing and Internalizing Symptoms Over 20 Years. *Developmental Psychology*, 41(5), 733-746. <https://doi.org/10.1037/0012-1649.41.5.733>
- Mathieu, R., Epinat-Duclos, J., Sigovan, M., Breton, A., Cheylus, A., Fayol, M., Thevenot, C., & Prado, J. (2018). What’s Behind a “+” Sign? Perceiving an Arithmetic Operator Recruits Brain Circuits for Spatial Orienting. *Cerebral Cortex*, 28(5), 1673-1684. <https://doi.org/10.1093/cercor/bhx064>
- May, T., Adesina, I., McGillivray, J., & Rinehart, N. J. (2019). Sex differences in neurodevelopmental disorders: *Current Opinion in Neurology*, 32(4), 622-626. <https://doi.org/10.1097/WCO.0000000000000714>
- Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers’ Precision of the Approximate Number System Predicts Later School Mathematics

- Performance. *PLoS ONE*, 6(9), e23749.
<https://doi.org/10.1371/journal.pone.0023749>
- McCann, J. C., & Ames, B. N. (2005). Is docosahexaenoic acid, an n-3 long-chain polyunsaturated fatty acid, required for development of normal brain function? An overview of evidence from cognitive and behavioral tests in humans and animals. *The American Journal of Clinical Nutrition*, 82(2), 281-295.
<https://doi.org/10.1093/ajcn.82.2.281>
- McCoach, D. B., & Siegle, D. (2003). The structure and function of academic self-concept in gifted and general education students. *Roeper Review*, 25(2), 61-65.
<https://doi.org/10.1080/02783190309554200>
- McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line : Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, 69(8), 1324-1333. <https://doi.org/10.3758/BF03192949>
- McEwan, P. J. (2003). Peer effects on student achievement : Evidence from Chile. *Economics of Education Review*, 22(2), 131-141. [https://doi.org/10.1016/S0272-7757\(02\)00005-5](https://doi.org/10.1016/S0272-7757(02)00005-5)
- Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012a). Phonological skills and their role in learning to read : A meta-analytic review. *Psychological Bulletin*, 138(2), 322-352. <https://doi.org/10.1037/a0026744>
- Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012b). Phonological skills and their role in learning to read : A meta-analytic review. *Psychological Bulletin*, 138(2), 322-352. <https://doi.org/10.1037/a0026744>
- Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the Home Learning Environment and Preschool Center Experience upon Literacy and Numeracy Development in Early Primary School. *Journal of Social Issues*, 64(1), 95-114. <https://doi.org/10.1111/j.1540-4560.2008.00550.x>
- Moilanen, K. L., Shaw, D. S., & Maxwell, K. L. (2010). Developmental cascades : Externalizing, internalizing, and academic competence from middle childhood to early adolescence. *Development and Psychopathology*, 22(3), 635-653.
<https://doi.org/10.1017/S0954579410000337>
- Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Streiftau, S., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Démonet, J.-F., Longeras, E., Valdois, S., George, F., ... Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, 29, 65-77.
<https://doi.org/10.1016/j.learninstruc.2013.09.003>
- Mortensen, E. L. (2002). The Association Between Duration of Breastfeeding and Adult Intelligence. *JAMA*, 287(18), 2365. <https://doi.org/10.1001/jama.287.18.2365>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Chetnut Hill, MA: Boston College.

- Muris, P., Meesters, C., & Diederens, R. (2005). Psychometric properties of the Big Five Questionnaire for Children (BFQ-C) in a Dutch sample of young adolescents. *Personality and Individual Differences, 38*(8), 1757-1769. <https://doi.org/10.1016/j.paid.2004.11.018>
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, Rimes, Vocabulary, and Grammatical Skills as Foundations of Early Reading Development : Evidence From a Longitudinal Study. *Developmental Psychology, 40*(5), 665-681. <https://doi.org/10.1037/0012-1649.40.5.665>
- National Early Literacy Panel. (2008). *Developing early literacy : Report of the National Early Literacy Panel*. <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Newcomb, A. F., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations : A meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychological Bulletin, 113*(1), 99-128. <https://doi.org/10.1037/0033-2909.113.1.99>
- Newman, T. M. (2008). Assessment of Giftedness in School-Age Children Using Measures of Intelligence or Cognitive Abilities. In S. I. Pfeiffer (Éd.), *Handbook of Giftedness in Children* (p. 161-176). Springer US. https://doi.org/10.1007/978-0-387-74401-8_9
- Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314-1334. <https://doi.org/10.1037/a0012702>
- Nguyen, T.-N.-N., Spencer-Smith, M., Zannino, D., Burnett, A., Scratch, S. E., Pascoe, L., Ellis, R., Cheong, J., Thompson, D., Inder, T., Doyle, L. W., & Anderson, P. J. (2018). Developmental Trajectory of Language From 2 to 13 Years in Children Born Very Preterm. *Pediatrics, 141*(5), e20172831. <https://doi.org/10.1542/peds.2017-2831>
- NICHD Early Child Care Research Network. (2005). Pathways to Reading : The Role of Oral Language in the Transition to Reading. *Developmental Psychology, 41*(2), 428-442. <https://doi.org/10.1037/0012-1649.41.2.428>
- Niklas, F., & Schneider, W. (2013). Home Literacy Environment and the beginning of reading and spelling. *Contemporary Educational Psychology, 38*(1), 40-50. <https://doi.org/10.1016/j.cedpsych.2012.10.001>
- Noël, M.-P. (2005). Finger gnosis : A predictor of numerical abilities in children? *Child Neuropsychology, 11*(5), 413-430. <https://doi.org/10.1080/09297040590951550>
- Nunes, G., Braga, L., Rossi, L., Lawisch, V., Nunes, L., & Dellatolas, G. (2008). Hand skill assessment with a reduced version of the Peg Moving Task (PMT-5) in children : Normative data and application in children with cerebral palsy. *Archives of Clinical Neuropsychology, 23*(1), 87-101. <https://doi.org/10.1016/j.acn.2007.08.007>

- OCDE. (2012). *Programme international pour le suivi des acquis des élèves (PISA)—Résultats du PISA 2012—Note par pays : France*.
<http://www.oecd.org/pisa/keyfindings/PISA-2012-results-france.pdf>
- O’Connell, M. (2018). The power of cognitive ability in explaining educational test performance, relative to other ostensible contenders. *Intelligence*, 66, 122-127.
<https://doi.org/10.1016/j.intell.2017.11.011>
- OECD. (2011). *Résultats du PISA 2009. Savoirs et savoir-faire des élèves : Performances des élèves en compréhension de l’écrit, en mathématiques et en sciences, Vol. I* (OECD Editions).
- OECD. (2015). *The ABC of Gender Equality in Education*. OECD Publishing.
<https://doi.org/10.1787/9789264229945-en>
- OECD. (2016a). *PISA 2015 Results (Volume I) : Excellence and Equity in Education*. OECD. <https://doi.org/10.1787/9789264266490-en>
- OECD. (2016b). *PISA 2015 Results (Volume II)*. OECD Publishing.
<https://doi.org/10.1787/9789264267510-en>
- OECD. (2019). *PISA 2018 Results (Volume I) : What Students Know and Can Do*. OECD. <https://doi.org/10.1787/5f07c754-en>
- Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children’s central conceptual structures in the domain of number. In *The role of central conceptual structures in the development of children’s thought : Monographs of the Society for Research in Child Development* (Vol. 1-2, p. 27-58). Blackwell Publishers.
- Olson, R. K., Keenan, J. M., Byrne, B., Samuelsson, S., Coventry, W. L., Corley, R., Wadsworth, S. J., Willcutt, E. G., Defries, J. C., Pennington, B. F., & Hulslander, J. (2011). Genetic and Environmental Influences on Vocabulary and Reading Development. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, 15(1), 26-46. <https://doi.org/10.1007/s11145-006-9018-x>
- O’Neil, H. F., Abedi, J., Miyoshi, J., & Mastergeorge, A. (2005). Monetary Incentives for Low-Stakes Tests. *Educational Assessment*, 10(3), 185-208.
https://doi.org/10.1207/s15326977ea1003_3
- Ors, E., Palomino, F., & Peyrache, E. (2013). Performance Gender Gap : Does Competition Matter? *Journal of Labor Economics*, 31(3), 443-499.
<https://doi.org/10.1086/669331>
- Ouellette, G. P. (2006). What’s meaning got to do with it : The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554-566. <https://doi.org/10.1037/0022-0663.98.3.554>
- Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability Differences Among People Who Have Commensurate Degrees Matter for Scientific Creativity. *Psychological Science*, 19(10), 957-961. <https://doi.org/10.1111/j.1467-9280.2008.02182.x>

- Passolunghi, M. C., Mammarella, I. C., & Altoè, G. (2008). Cognitive Abilities as Precursors of the Early Acquisition of Mathematical Skills During First Through Second Grades. *Developmental Neuropsychology*, 33(3), 229-250. <https://doi.org/10.1080/87565640801982320>
- Pastorelli, C., Caprara, G. V., Barbaranelli, C., Rola, J., Rozsa, S., & Bandura, A. (2001). The Structure of Children's Perceived Self-Efficacy : A Cross-National Study. *European Journal of Psychological Assessment*, 17(2), 87-97. <https://doi.org/10.1027//1015-5759.17.2.87>
- Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory : Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, 108(4), 455-473. <https://doi.org/10.1037/edu0000079>
- Penner-Wilger, M., Fast, L., LeFevre, J., Smith-Chant, B., Skwarchuk, S., Kamawar, D., & Bisanz, J. (2007). *The Foundations of Numeracy : Subitizing, Finger Gnosia, and Fine Motor Ability*. 1385-1390.
- Peters, M. P., & Bain, S. K. (2011). Bullying and Victimization Rates among Gifted and High-Achieving Students. *Journal for the Education of the Gifted*, 34(4), 624-643. <https://doi.org/10.1177/016235321103400405>
- Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., Pettit, G. S., & Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology*, 122(2), 542-557. <https://doi.org/10.1037/a0031963>
- Peterson, J. S., & Ray, K. E. (2006). Bullying and the Gifted : Victims, Perpetrators, Prevalence, and Effects. *Gifted Child Quarterly*, 50(2), 148-168. <https://doi.org/10.1177/001698620605000206>
- Peyre, H., Bernard, J. Y., Hoertel, N., Forhan, A., Charles, M.-A., De Agostini, M., Heude, B., & Ramus, F. (2016). Differential effects of factors influencing cognitive development at the age of 5-to-6 years. *Cognitive Development*, 40, 152-162. <https://doi.org/10.1016/j.cogdev.2016.10.001>
- Peyre, H., Galera, C., van der Waerden, J., Hoertel, N., Bernard, J. Y., Melchior, M., & Ramus, F. (2016). Relationship between early language skills and the development of inattention/hyperactivity symptoms during the preschool period : Results of the EDEN mother-child cohort. *BMC Psychiatry*, 16(1). <https://doi.org/10.1186/s12888-016-1091-3>
- Peyre, H., Ramus, F., Melchior, M., Forhan, A., Heude, B., & Gauvrit, N. (2016). Emotional, behavioral and social difficulties among high-IQ children during the preschool period : Results of the EDEN mother-child cohort. *Personality and Individual Differences*, 94, 366-371. <https://doi.org/10.1016/j.paid.2016.02.014>
- Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *School Psychology Review*, 33, 444-458.

- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33-41. <https://doi.org/10.1016/j.cognition.2010.03.012>
- Pingault, J.-B., O'Reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijdsdijk, F., & Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*, *19*(9), 566-580. <https://doi.org/10.1038/s41576-018-0020-3>
- Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition*, *109*(3), 408-415. <https://doi.org/10.1016/j.cognition.2008.09.003>
- Pinheiro-Chagas, P., Didino, D., Haase, V. G., Wood, G., & Knops, A. (2018). The Developmental Trajectory of the Operational Momentum Effect. *Frontiers in Psychology*, *9*. <https://doi.org/10.3389/fpsyg.2018.01062>
- Pinheiro-Chagas, P., Dotan, D., Piazza, M., & Dehaene, S. (2017). Finger Tracking Reveals the Covert Stages of Mental Arithmetic. *Open Mind*, *1*(1), 30-41. https://doi.org/10.1162/OPMI_a_00003
- Piquard-Kipffer, A., & Sprenger-Charolles, L. (2013). Early predictors of future reading skills : A follow-up of French-speaking children from the beginning of kindergarten to the end of the second grade (age 5 to 8). *L'Année Psychologique*, *Vol. 113*(4), 491-521.
- Pitchford, N. J., Papini, C., Outhwaite, L. A., & Gulliford, A. (2016). Fine Motor Skills Predict Maths Ability Better than They Predict Reading Ability in the Early Primary School Years. *Frontiers in Psychology*, *7*, 783. <https://doi.org/10.3389/fpsyg.2016.00783>
- Plaza, M. (2001). The interaction between phonological processing, syntactic awareness and reading : A longitudinal study from Kindergarten to Grade 1: *First Language*. <https://doi.org/10.1177/014272370102106101>
- Plaza, M., & Cohen, H. (2007). The contribution of phonological awareness and visual attention in early reading and spelling. *Dyslexia (Chichester, England)*, *13*(1), 67-76. <https://doi.org/10.1002/dys.330>
- Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, Nurture, and Cognitive Development from 1 to 16 Years : A Parent-Offspring Adoption Study. *Psychological Science*, *8*(6), 442-447. <https://doi.org/10.1111/j.1467-9280.1997.tb00458.x>
- Plomin, R., Owen, M., & McGuffin, P. (1994). The genetic basis of complex human behaviors. *Science*, *264*(5166), 1733-1739. <https://doi.org/10.1126/science.8209254>
- Plomin, Robert, & Bergeman, C. S. (1991). The nature of nurture : Genetic influence on "environmental" measures. *Behavioral and Brain Sciences*, *14*(3), 373-386. <https://doi.org/10.1017/S0140525X00070278>

- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322-338. <https://doi.org/10.1037/a0014996>
- Postlethwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains : A meta-analysis* [Doctor of Philosophy, University of Iowa]. <https://doi.org/10.17077/etd.zopi8wvs>
- Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A. S., Minas, J. E., & Booth, J. R. (2011). Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. *Human Brain Mapping*, *32*(11), 1932-1947. <https://doi.org/10.1002/hbm.21159>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891. <https://doi.org/10.3758/BRM.40.3.879>
- Protivínský, T., & Münich, D. (2018). Gender Bias in teachers' grading : What is in the grade. *Studies in Educational Evaluation*, *59*, 141-149. <https://doi.org/10.1016/j.stueduc.2018.07.006>
- Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The Components of the Simple View of Reading : A Confirmatory Factor Analysis. *Reading Psychology*, *33*(3), 217-240. <https://doi.org/10.1080/02702711.2010.507626>
- Quillet, L. (2012, novembre 23). *70% des surdoués sont en échec scolaire*. Le Figaro Etudiant. <http://etudiant.lefigaro.fr/les-news/actu/detail/article/70-des-surdoues-sont-en-echec-scolaire-554/>
- Rabiner, D., & Coie, J. D. (2000). Early Attention Problems and Children's Reading Achievement : A Longitudinal Investigation. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(7), 859-867. <https://doi.org/10.1097/00004583-200007000-00014>
- Radloff, L. S. (1977). The CES-D Scale : A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, *1*(3), 385-401. <https://doi.org/10.1177/014662167700100306>
- Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S. C., Ellis, C., Shakeshaft, C., Seghier, M. L., & Price, C. J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature*, *479*(7371), 113-116. <https://doi.org/10.1038/nature10514>
- Ramus, F. (2003). Developmental dyslexia : Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212-218. [https://doi.org/10.1016/s0959-4388\(03\)00035-7](https://doi.org/10.1016/s0959-4388(03)00035-7)
- Rapport, M. D., Scanlan, S. W., & Denney, C. B. (1999). Attention-Deficit/Hyperactivity Disorder and Scholastic Achievement : A Model of Dual Developmental Pathways. *Journal of Child Psychology and Psychiatry*, *40*(8), 1169-1183. <https://doi.org/10.1111/1469-7610.00534>

- Raven, J. C. (1998). *Raven's Progressive Matrices*. Oxford Psychologists Press, Ltd.
- Razaz, N., Boyce, W. T., Brownell, M., Jutte, D., Tremlett, H., Marrie, R. A., & Joseph, K. S. (2016). Five-minute Apgar score as a marker for developmental vulnerability at 5 years of age. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *101*(2), F114-F120. <https://doi.org/10.1136/archdischild-2015-308458>
- Reardon, S. F., Kalogrides, D., Fahle, E. M., Podolsky, A., & Zárate, R. C. (2018). The Relationship Between Test Item Format and Gender Achievement Gaps on Math and ELA Tests in Fourth and Eighth Grades. *Educational Researcher*, *47*(5), 284-294. <https://doi.org/10.3102/0013189X18762105>
- Reichenberg, A., Gross, R., Weiser, M., Bresnahan, M., Silverman, J., Harlap, S., Rabinowitz, J., Shulman, C., Malaspina, D., Lubin, G., Knobler, H. Y., Davidson, M., & Susser, E. (2006). Advancing Paternal Age and Autism. *Archives of General Psychiatry*, *63*(9), 1026. <https://doi.org/10.1001/archpsyc.63.9.1026>
- Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement : A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology*, *107*(3), 645-662. <https://doi.org/10.1037/edu0000012>
- Reilly, D., Neumann, D. L., & Andrews, G. (2018). Gender differences in reading and writing achievement : Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*. <https://doi.org/10.1037/amp0000356>
- Reis, S. M., & McCoach, D. B. (2000). The Underachievement of Gifted Students : What Do We Know and Where Do We Go? *Gifted Child Quarterly*, *44*(3), 152-170. <https://doi.org/10.1177/001698620004400302>
- Reuhkala, M. (2001). Mathematical Skills in Ninth-graders : Relationship with visuo-spatial abilities and working memory. *Educational Psychology*, *21*(4), 387-399. <https://doi.org/10.1080/01443410120090786>
- Rhoades, B. L., Warren, H. K., Domitrovich, C. E., & Greenberg, M. T. (2011). Examining the link between preschool social-emotional competence and first grade academic achievement : The role of attention skills. *Early Childhood Research Quarterly*, *26*(2), 182-191. <https://doi.org/10.1016/j.ecresq.2010.07.003>
- Riggs, N. R., Jahromi, L. B., Razza, R. P., Dillworth-Bart, J. E., & Mueller, U. (2006). Executive function and the promotion of social-emotional competence. *Journal of Applied Developmental Psychology*, *27*(4), 300-309. <https://doi.org/10.1016/j.appdev.2006.04.002>
- Ritchie, S. J., & Bates, T. C. (2013). Enduring Links From Childhood Mathematics and Reading Achievement to Adult Socioeconomic Status. *Psychological Science*, *24*(7), 1301-1308. <https://doi.org/10.1177/0956797612466268>
- Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychological Science*, *29*(8), 1358-1369. <https://doi.org/10.1177/0956797618774253>

- Robertson, E. (1991). Neglected Dropouts The Gifted and Talented. *Equity & Excellence in Education*, 25(1), 62-73. <https://doi.org/10.1080/1066568910250112>
- Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the Threshold Hypothesis : Even Among the Gifted and Top Math/Science Graduate Students, Cognitive Abilities, Vocational Interests, and Lifestyle Preferences Matter for Career Choice, Performance, and Persistence. *Current Directions in Psychological Science*, 19(6), 346-351. <https://doi.org/10.1177/0963721410391442>
- Robinson, J. P., & Lubienski, S. T. (2011). The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School : Examining Direct Cognitive Assessments and Teacher Ratings. *American Educational Research Journal*, 48(2), 268-302. <https://doi.org/10.3102/0002831210372249>
- Rocher, T. (2008). *Note d'information 08.38—Lire, écrire, compter : Les performances des élèves de CM2 à vingt ans d'intervalle 1987-2007*. <http://www.cafepedagogique.net/lexpresso/Documents/docsjoints/NI-0838.pdf>
- Roeser, R., Eccles, J., & Strobel, K. (1998). Linking the Study of Schooling and Mental Health : Selected Issues and Empirical Illustrations at the Level of the Individual. *Educational Psychologist*, 33, 153-176.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35(1), 83-92. <https://doi.org/10.1016/j.intell.2006.05.004>
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades : A meta-analysis. *Intelligence*, 53, 118-137. <https://doi.org/10.1016/j.intell.2015.09.002>
- Rousseau, S., & Tardieu, F. (2004). *La macro SAS CUBE d'échantillonnage équilibré. Documentation de l'utilisateur*.
- Routitsky, A., & Turner, R. (2003). *Item format types and their influence on cross-national comparisons of student performance*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Roznowski, M., Hong, S., & Reith, J. (2000). A Further Look at Youth Intellectual Giftedness and its Correlates : Values, Interests, Performance, and Behavior. *Intelligence*, 28(2), 87-113. [https://doi.org/10.1016/S0160-2896\(99\)00032-X](https://doi.org/10.1016/S0160-2896(99)00032-X)
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.
- Rutter, M. (2007). Gene-environment interdependence. *Developmental Science*, 10(1), 12-18. <https://doi.org/10.1111/j.1467-7687.2007.00557.x>
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization : Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57(5), 749-761. <https://doi.org/10.1037/0022-3514.57.5.749>

- Saha, S., Barnett, A. G., Foldi, C., Burne, T. H., Eyles, D. W., Buka, S. L., & McGrath, J. J. (2009). Advanced Paternal Age Is Associated with Impaired Neurocognitive Outcomes during Infancy and Childhood. *PLoS Medicine*, *6*(3), e1000040. <https://doi.org/10.1371/journal.pmed.1000040>
- Sammons, P., Toth, K., Sylva, K., Melhuish, E., Siraj, I., & Taggart, B. (2015). The long-term role of the home learning environment in shaping students' academic attainment in secondary school. *Journal of Children's Services*, *10*(3), 189-201. <https://doi.org/10.1108/JCS-02-2015-0007>
- Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., & Reichenberg, A. (2017). The Heritability of Autism Spectrum Disorder. *JAMA*, *318*(12), 1182. <https://doi.org/10.1001/jama.2017.12141>
- Sandrini, M., Miozzo, A., Cotelli, M., & Cappa, S. F. (2003). The Residual Calculation Abilities of a Patient with Severe Aphasia : Evidence for a Selective Deficit of Subtraction Procedures. *Cortex*, *39*(1), 85-96. [https://doi.org/10.1016/S0010-9452\(08\)70076-5](https://doi.org/10.1016/S0010-9452(08)70076-5)
- Santos, D., & Primi, R. (2014). *Social and emotional development and school learning : A measurement proposal in support of public policy*. Instituto Ayrton Senna. <http://educacaosec21.org.br/wp-content/uploads/2013/07/Social-and-emotional-developmente-and-school-learning.pdf>
- Savage, R., Pillay, V., & Melidona, S. (2008). Rapid Serial Naming Is a Unique Predictor of Spelling in Children. *Journal of Learning Disabilities*, *41*(3), 235-250. <https://doi.org/10.1177/0022219408315814>
- Scaff, C., & Cristia, A. (in prep.). Socio-economic status correlates with infants' word comprehension : A meta-analysis.
- Scheiber, C., Reynolds, M. R., Hajovsky, D. B., & Kaufman, A. S. (2015). Gender Differences in Achievement in a Large, Nationally Representative Sample of Children and Adolescents : Gender and Achievement. *Psychology in the Schools*, *52*(4), 335-348. <https://doi.org/10.1002/pits.21827>
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In *Contemporary intellectual assessment : Theories, tests, and issues* (The Guilford Press, p. 99-144).
- Seigneuric, A., & Ehrlich, M.-F. (2005). Contribution of Working Memory Capacity to Children's Reading Comprehension : A Longitudinal Investigation. *Reading and Writing*, *18*(7-9), 617-656. <https://doi.org/10.1007/s11145-005-2038-0>
- Senechal, M. (2006). Testing the Home Literacy Model : Parent Involvement in Kindergarten Is Differentially Related to Grade 4 Reading Comprehension, Fluency, Spelling, and Reading for Pleasure. *Scientific Studies of Reading*, *10*(1), 59-87. https://doi.org/10.1207/s1532799xssr1001_4
- Senechal, M., & LeFevre, J.-A. (2002). Parental Involvement in the Development of Children's Reading Skill : A Five-Year Longitudinal Study. *Child Development*, *73*(2), 445-460. <https://doi.org/10.1111/1467-8624.00417>

- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology (London, England: 1953)*, 94(Pt 2), 143-174. <https://doi.org/10.1348/000712603321661859>
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice : The perils of overreliance on an « outlier » orthography. *Psychological Bulletin*, 134(4), 584-615. <https://doi.org/10.1037/0033-2909.134.4.584>
- Shojaei, T., Wazana, A., Pitrou, I., & Kovess, V. (2009). The strengths and difficulties questionnaire : Validation study in French school-aged children and cross-cultural comparisons. *Social Psychiatry and Psychiatric Epidemiology*, 44(9), 740-747. <https://doi.org/10.1007/s00127-008-0489-8>
- Siegler, R. S., & Booth, J. L. (2004). Development of Numerical Estimation in Young Children. *Child Development*, 75(2), 428-444. <https://doi.org/10.1111/j.1467-8624.2004.00684.x>
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early Predictors of High School Mathematics Achievement. *Psychological Science*, 23(7), 691-697. <https://doi.org/10.1177/0956797612440101>
- Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction : How do children know what to do? In *Origins of cognitive skills* (Erlbaum, p. 229-293).
- Silverman, I. (2003). Gender Differences in Delay of Gratification : A Meta-Analysis. *Sex Roles*, 49(9), 451-463. <https://doi.org/10.1023/A:1025872421115>
- Simmons, F. R., & Singleton, C. (2008). Do weak phonological representations impact on arithmetic development? A review of research into arithmetic and dyslexia. *Dyslexia*, 14(2), 77-94. <https://doi.org/10.1002/dys.341>
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement : A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417-453. <https://doi.org/10.3102/00346543075003417>
- Sirin, Selcuk R. (2005). Socioeconomic Status and Academic Achievement : A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417-453. <https://doi.org/10.3102/00346543075003417>
- Slusser, E., Ribner, A., & Shusterman, A. (2019). Language counts : Early language mediates the relationship between parent education and children's math ability. *Developmental Science*, 22(3). <https://doi.org/10.1111/desc.12773>
- Social & Emotional Issues / National Association for Gifted Children.* (s. d.). Consulté 16 juillet 2018, à l'adresse <https://www.nagc.org/resources-publications/resources-parents/social-emotional-issues>
- Sohr-Preston, S. L., & Scaramella, L. V. (2006). Implications of Timing of Maternal Depressive Symptoms for Early Cognitive and Language Development. *Clinical Child and Family Psychology Review*, 9(1), 65-83. <https://doi.org/10.1007/s10567-006-0004-2>

- Spearman, C. (1904). « General Intelligence » Objectively Determined and Measured. *The American Journal of Psychology*.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. <https://doi.org/10.1006/jesp.1998.1373>
- Sprietsma, M. (2013). Discrimination in grading : Experimental evidence from primary school teachers. *Empirical Economics*, 45(1), 523-538. <https://doi.org/10.1007/s00181-012-0609-x>
- Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, 110(45), 18116-18120. <https://doi.org/10.1073/pnas.1302751110>
- Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The Multiple Subfunctions of Attention : Differential Developmental Gateways to Literacy and Numeracy: **Attention Subfunctions as Gateways**. *Child Development*, 83(6), 2028-2041. <https://doi.org/10.1111/j.1467-8624.2012.01809.x>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.
- Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement : What are the roles of personality and achievement motivation? *European Journal of Personality*, 22(3), 185-209. <https://doi.org/10.1002/per.676>
- Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with Higher Levels of Gender Equality Show Larger National Sex Differences in Mathematics Anxiety and Relatively Lower Parental Mathematics Valuation for Girls. *PLOS ONE*, 11(4), e0153857. <https://doi.org/10.1371/journal.pone.0153857>
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading : Evidence from a longitudinal structural model. *Developmental Psychology*, 38(6), 934-947.
- Streissguth, A. P., Barr, H. M., Olson, H. C., Sampson, P. D., Bookstein, F. L., & Burgess, D. M. (1994). Drinking during pregnancy decreases word attack and arithmetic scores on standardized tests : Adolescent data from a population-based prospective study. *Alcoholism, Clinical and Experimental Research*, 18(2), 248-254. <https://doi.org/10.1111/j.1530-0277.1994.tb00009.x>
- Su, M., Peyre, H., Song, S., McBride, C., Tardif, T., Li, H., Zhang, Y., Liang, W., Zhang, Z., Ramus, F., & Shu, H. (2017). The influence of early linguistic skills and family factors on literacy acquisition in Chinese children : Follow-up from age 3 to age 11. *Learning and Instruction*, 49, 54-63. <https://doi.org/10.1016/j.learninstruc.2016.12.003>
- Suárez-Pellicioni, M., Fuchs, L., & Booth, J. R. (2019). Temporo-frontal activation during phonological processing predicts gains in arithmetic facts in young

- children. *Developmental Cognitive Neuroscience*, 40, 100735.
<https://doi.org/10.1016/j.dcn.2019.100735>
- Suggate, S., Pufke, E., & Stoeger, H. (2018). Do fine motor skills contribute to early reading development? *Journal of Research in Reading*, 41(1), 1-19.
<https://doi.org/10.1111/1467-9817.12081>
- Suggate, S., Pufke, E., & Stoeger, H. (2019). Children's fine motor skills in kindergarten predict reading in grade 1. *Early Childhood Research Quarterly*, 47, 248-258.
<https://doi.org/10.1016/j.ecresq.2018.12.015>
- Talsma, K., Schütz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa) : A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, 61, 136-150.
<https://doi.org/10.1016/j.lindif.2017.11.015>
- Tam, V. C., & Chan, R. M. (2009). Parental involvement in primary children's homework in Hong Kong. *The School Community Journal*, 19(2), 81-100.
- Taylor, H. G., Espy, K. A., & Anderson, P. J. (2009). Mathematics deficiencies in children with very low birth weight or very preterm birth. *Developmental Disabilities Research Reviews*, 15(1), 52-59. <https://doi.org/10.1002/ddrr.51>
- Terman, L. M. (1926a). *Genetic Studies Of Genius, Volume I: Mental And Physical Traits Of A Thousand Gifted Children*. Stanford University Press.
- Terman, L. M. (1926b). School Progress and Educational History. In *Genetic Studies Of Genius, Volume I: Mental And Physical Traits Of A Thousand Gifted Children* (p. 253-287). Stanford University Press.
- Terman, L. M. (1926c). Tests of School Accomplishment and of General Information. In *Genetic Studies Of Genius, Volume I: Mental And Physical Traits Of A Thousand Gifted Children* (p. 289-306). Stanford University Press.
- Terrassier, J.-C. (2009). *Les enfants surdoués ou La précocité embarrassante*. Esf Editeur.
- Terrier, C. (2015). Giving a Little Help to Girls ? Evidence on Grade Discrimination and its Effect on Students' Achievement. *Centre for Economic Performance London School of Economics and Political Science, CEP Discussion Paper No 1341*.
- Terriot, K. (2014). Testons les tests : Le RCC (Raisonnement sur Cartes de Chartier). *ANAE - Approche Neuropsychologique des Apprentissages chez l'Enfant*, 26(129), 179-183.
- Testa, M. (2003). THE EFFECTS OF PRENATAL ALCOHOL EXPOSURE ON INFANT MENTAL DEVELOPMENT : A META-ANALYTICAL REVIEW. *Alcohol and Alcoholism*, 38(4), 295-304. <https://doi.org/10.1093/alcalc/agg087>
- Torppa, M., Georgiou, G., Lerkkanen, M.-K., Niemi, P., Poikkeus, & Nurmi, J.-E. (2016). Examining the Simple View of Reading in a Transparent Orthography : A

- Longitudinal Study From Kindergarten to Grade 3. *Merrill-Palmer Quarterly*, 62, 179. <https://doi.org/10.13110/merrpalmquar1982.62.2.0179>
- Tosto, M. G., Hayiou-Thomas, M. E., Harlaar, N., Prom-Wormley, E., Dale, P. S., & Plomin, R. (2017). The genetic architecture of oral language, reading fluency, and reading comprehension : A twin study from 7 to 16 years. *Developmental Psychology*, 53(6), 1115-1129. <https://doi.org/10.1037/dev0000297>
- Träff, U., Olsson, L., Skagerlund, K., & Östergren, R. (2018). Cognitive mechanisms underlying third graders' arithmetic skills : Expanding the pathways to mathematics model. *Journal of Experimental Child Psychology*, 167, 369-387. <https://doi.org/10.1016/j.jecp.2017.11.010>
- Trosseille, B., Champault, F., & Lieury, A. (2013). Évaluation de 30 000 élèves de 6^e du collège français. Présentation et introduction. *Bulletin de psychologie*, Numéro 523(1), 3. <https://doi.org/10.3917/bupsy.523.0003>
- Trzaskowski, M., Eley, T. C., Davis, O. S. P., Doherty, S. J., Hanscombe, K. B., Meaburn, E. L., Haworth, C. M. A., Price, T., & Plomin, R. (2013). First Genome-Wide Association Study on Anxiety-Related Behaviours in Childhood. *PLoS ONE*, 8(4), e58676. <https://doi.org/10.1371/journal.pone.0058676>
- Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P. S., & Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, 42, 83-88. <https://doi.org/10.1016/j.intell.2013.11.002>
- Tseng, P.-T., Yen, C.-F., Chen, Y.-W., Stubbs, B., Carvalho, A. F., Whiteley, P., Chu, C.-S., Li, D.-J., Chen, T.-Y., Yang, W.-C., Tang, C.-H., Liang, H.-Y., Yang, W.-C., Wu, C.-K., & Lin, P.-Y. (2019). Maternal breastfeeding and attention-deficit/hyperactivity disorder in children : A meta-analysis. *European Child & Adolescent Psychiatry*, 28(1), 19-30. <https://doi.org/10.1007/s00787-018-1182-4>
- Twilhaar, E. S., Wade, R. M., de Kieviet, J. F., van Goudoever, J. B., van Elburg, R. M., & Oosterlaan, J. (2018). Cognitive Outcomes of Children Born Extremely or Very Preterm Since the 1990s and Associated Risk Factors : A Meta-analysis and Meta-regression. *JAMA Pediatrics*, 172(4), 361. <https://doi.org/10.1001/jamapediatrics.2017.5323>
- Uittenhove, K., Thevenot, C., & Barrouillet, P. (2016). Fast automated counting procedures in addition problem solving : When are they used and why are they mistaken for retrieval? *Cognition*, 146, 289-303. <https://doi.org/10.1016/j.cognition.2015.10.008>
- Valdois, S., Roulin, J.-L., & Line Bosse, M. (2019). Visual attention modulates reading acquisition. *Vision Research*, 165, 152-161. <https://doi.org/10.1016/j.visres.2019.10.011>
- van der Ven, S. H. G., van der Maas, H. L. J., Straatemeier, M., & Jansen, B. R. J. (2013). Visuospatial working memory and mathematical ability at different ages

- throughout primary school. *Learning and Individual Differences*, 27, 182-192.
<https://doi.org/10.1016/j.lindif.2013.09.003>
- van Harskamp, N. J., & Cipolotti, L. (2001). Selective Impairments for Addition, Subtraction and Multiplication. Implications for the Organisation of Arithmetical Facts. *Cortex*, 37(3), 363-388. [https://doi.org/10.1016/S0010-9452\(08\)70579-3](https://doi.org/10.1016/S0010-9452(08)70579-3)
- VanderWeele, T. J. (2016). Mediation Analysis : A Practitioner's Guide. *Annual Review of Public Health*, 37(1), 17-32. <https://doi.org/10.1146/annurev-publhealth-032315-021402>
- Vannetzel, L. (2009). « Mon enfant est-il surdoué ? » Que demande-t-on ? *Pratiques Psychologiques*, 15(3), 327-341. <https://doi.org/10.1016/j.prps.2008.07.003>
- Verguts, T., & Fias, W. (2005). Interacting neighbors : A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition*, 33(1), 1-16.
<https://doi.org/10.3758/BF03195293>
- Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia : A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences*, 14(2), 57-63.
<https://doi.org/10.1016/j.tics.2009.12.003>
- von Stumm, S. (2017). Socioeconomic status amplifies the achievement gap throughout compulsory education independent of intelligence. *Intelligence*, 60, 57-62.
<https://doi.org/10.1016/j.intell.2016.11.006>
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities : A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 2, 250-270.
- Voyer, Daniel, & Voyer, S. D. (2014). Gender differences in scholastic achievement : A meta-analysis. *Psychological Bulletin*, 140(4), 1174-1204.
<https://doi.org/10.1037/a0036620>
- Walsh, J. J., Barnes, J. D., Cameron, J. D., Goldfield, G. S., Chaput, J.-P., Gunnell, K. E., Ledoux, A.-A., Zemek, R. L., & Tremblay, M. S. (2018). Associations between 24 hour movement behaviours and global cognition in US children : A cross-sectional observational study. *The Lancet Child & Adolescent Health*, 2(11), 783-791. [https://doi.org/10.1016/S2352-4642\(18\)30278-5](https://doi.org/10.1016/S2352-4642(18)30278-5)
- Watkins, M. W., & Smith, L. G. (2013). Long-term stability of the Wechsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment*, 25(2), 477-483.
<https://doi.org/10.1037/a0031653>
- Webb, J. T. (2014, juin 2). *Gifted Children and Adults—Neglected Areas of Practice*. National Register. <https://www.nationalregister.org/pub/the-national-register-report-pub/spring-2014-issue/gifted-children-and-adults-neglected-areas-of-practice/>
- Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence* (3rd ed.). The Psychological Corporation.

- Wechsler, D. (2003). *Wechsler intelligence scale for children—Fourth Edition (WISC-IV)* (4th ed.). The Psychological Corporation.
- Wechsler, D. (2004). *WPPSI-III, Troisième édition de l’Echelle d’Intelligence de Wechsler pour la période préscolaire et primaire*. ECPA (Editions du Centre de Psychologie Appliquée).
- Wechsler, D. (2005). *Échelle d’intelligence de Wechsler pour enfants : WISC-IV*. ECPA (Editions du Centre de Psychologie Appliquée).
- Whitehouse, A. J. O., Robinson, M., Li, J., & Oddy, W. H. (2011). Duration of breast feeding and language ability in middle childhood. *Paediatric and Perinatal Epidemiology*, 25(1), 44-52. <https://doi.org/10.1111/j.1365-3016.2010.01161.x>
- Whitehurst, G. J., & Lonigan, C. J. (1998). Child Development and Emergent Literacy. *Child Development*, 69(3), 848-872. <https://doi.org/10.1111/j.1467-8624.1998.tb06247.x>
- Willcutt, E. G. (2012). The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder : A Meta-Analytic Review. *Neurotherapeutics*, 9(3), 490-499. <https://doi.org/10.1007/s13311-012-0135-8>
- Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric Comorbidity in Children and Adolescents with Reading Disability. *Journal of Child Psychology and Psychiatry*, 41(8), 1039-1048. <https://doi.org/10.1111/1469-7610.00691>
- Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Lawrence Erlbaum Associates Publishers.
- Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math Anxiety in Second and Third Graders and Its Relation to Mathematics Achievement. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00162>
- Wu, S. S., Willcutt, E. G., Escovar, E., & Menon, V. (2014). Mathematics Achievement and Anxiety and Their Relation to Internalizing and Externalizing Behaviors. *Journal of Learning Disabilities*, 47(6), 503-514. <https://doi.org/10.1177/0022219412473154>
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA : A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*, 88(1), 76-82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5), 678-704. <https://doi.org/10.1080/01443410.2018.1546831>
- Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence*, 35(5), 489-501. <https://doi.org/10.1016/j.intell.2007.02.003>

- Zhang, X., Koponen, T., Räsänen, P., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2014). Linguistic and Spatial Skills Predict Early Arithmetic Development via Counting Sequence Knowledge. *Child Development, 85*(3), 1091-1107. <https://doi.org/10.1111/cdev.12173>
- Zhang, X., Räsänen, P., Koponen, T., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2017). Knowing, applying, and reasoning about arithmetic : Roles of domain-general and numerical skills in multiple domains of arithmetic learning. *Developmental Psychology, 53*(12), 2304-2318. <https://doi.org/10.1037/dev0000432>
- Zhang, Y., Tardif, T., Shu, H., Li, H., Liu, H., McBride-Chang, C., Liang, W., & Zhang, Z. (2013). Phonological skills and vocabulary knowledge mediate socioeconomic status effects in predicting reading outcomes for Chinese children. *Developmental Psychology, 49*(4), 665-671. <https://doi.org/10.1037/a0028612>
- Zhou, X., Chen, C., Zang, Y., Dong, Q., Chen, C., Qiao, S., & Gong, Q. (2007). Dissociated brain organization for single-digit addition and multiplication. *NeuroImage, 35*(2), 871-880. <https://doi.org/10.1016/j.neuroimage.2006.12.017>
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Fásca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic Depth and Its Impact on Universal Predictors of Reading : A Cross-Language Investigation. *Psychological Science, 21*(4), 551-559. <https://doi.org/10.1177/0956797610363406>
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers, 28*(4), 504-515. <https://doi.org/10.3758/BF03200539>
- Zimmerman, B. J. (2000). Self-Efficacy : An Essential Motive to Learn. *Contemporary Educational Psychology, 25*(1), 82-91. <https://doi.org/10.1006/ceps.1999.1016>
- Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes : Evidence from a Natural Experiment. *Review of Economics and Statistics, 85*(1), 9-23. <https://doi.org/10.1162/003465303762687677>