LONDON'S GLOBAL UNIVERSITY

Alexis Karamanos

Attainment grouping in primary school, verbal skills, and psychological symptoms in the United Kingdom; Evidence from the Millennium Cohort Study

Faculty of Population Health Sciences

Institute of Epidemiology and Health Care

UCL Research Department of Epidemiology and Public Health

Thesis submitted for the degree of Ph.D

2020

Declaration

I, Alexis Karamanos confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

This work has been funded by an Economic and Social Research Council (ESRC) Studentship.

Abstract

Background

There is a lack of studies on the role of primary school attainment grouping on children's verbal skills and psychological symptoms (externalising and internalising) in a contemporary United Kingdom (UK) context. Socioeconomic inequalities in cognitive and psychological outcomes increase during compulsory education in the UK. However, no UK studies exist on the contributing role of attainment grouping in primary school. This PhD examines associations between attainment grouping, family income, initial verbal differences (age 5), and the development of verbal skills and psychological symptoms.

Methods

By using the Millennium Cohort Study (MCS), fixed effects difference-in-difference modelling examined associations between attainment grouping transition, and the change in verbal skills and psychological symptoms from 7 to 11 years. Growth curve modelling examined associations between attainment grouping at age 7 and trajectories of psychological symptoms from 7 to 14 years.

Results

Overall, no association was observed between attainment grouping transition and a change in MCS participants' verbal skills or psychological symptoms from ages 7 to 11 years. However, contingent on verbal skills at age 5 (high vs low), the gap in verbal skills grew over time for children who transitioned from mixed attainment classes to attainment groups compared with those who stayed in mixed attainment classes throughout.

Overall, attainment grouping at age 7 was not associated with trajectories of psychological symptoms from age 7 to age 14 and was not associated with an increase in socioeconomic inequalities in psychological symptoms over time. Attainment grouping at age 7 was associated with a slightly slower increase in internalising symptoms over time among those children who entered primary school with lower verbal skills.

Conclusions

Overall, attainment grouping in primary school was not associated with a change in verbal skills or psychological symptoms. Attainment grouping transition was associated with an increasing gap in verbal skills during primary school.

Acknowledgements

This PhD would not have been possible without the support of several people. First, I would like to thank my primary supervisor Yvonne Kelly for inspiring me with every conversation, encouraging me to see the wider context of research and supporting me to grow in confidence throughout this process. I would like to thank Anne McMunn, Anja Heilmann and Seeromanie Harding for sharing their extraordinary, support and guidance over the course of my PhD. I thank David Blane for his valuable support, advice, and mentoring. I would like to thank the Onassis Foundation for its financial support and the ESRC for providing funding for this research, the children of the Millennium Cohort Study for their participation and everyone who made possible the collection of these data. During my time as a PhD student at UCL, I have been very fortunate to be within a group of people who have supported one another through issues and celebrated one another's successes, thank you all. I would like to express my dearest thanks to my parents, for their resolute belief that I can achieve great things in life with persistence and effort. I need to thank a number of other important people who have had a hand in encouraging me along this journey: a great friend and mentor, Evangelos Drakopoulos; my brother Lazaros for being by my side through every step of my life. Finally, I owe a special thanks to my life partner and best friend Elena, for her enduring love, support, and encouragement, for being my home and safe place.

Impact statement

This PhD contributes to the limited body of research on the consequences of attainment grouping (streaming or setting) during primary school in a UK contemporary context is currently lacking. This PhD attempted to overcome methodological issues of previous academic research such as unobserved time-invariant and time-varying confounding bias. By applying fixed effects difference-in-difference modelling, it was possible to account for hard to measure differences between children, that do not change over time, and shared changes over time between primary schools in England and Wales e.g external factors/shocks such as the Great Recession and its aftermath (2008-2012). In addition, this PhD is the first to provide evidence on the long-term association between early attainment grouping (age seven) and trajectories of psychological symptoms from 7 to 14 years via growth curve modelling.

The transition to attainment grouping between ages 7 and 11 years was not associated with an overall increase in verbal cognitive attainment. Nevertheless, attainment grouping transition had a disproportionate effect on low attaining students. Low attaining students hardly improved their verbal skills compared to high-attaining students who improved their verbal skills considerably. Therefore, the gap in verbal skills at school entry (high versus low scores) widened for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups.

Abolishing between-class attainment grouping in the UK is difficult. Given the reluctance of the UK schools to adopt mixed attainment methods and an increased level of parental choice, attention could therefore be focused on applying more equitable principles in attainment grouping with the aim of improving the quality of pedagogy for children who enter primary school with a lower level of verbal skills. If primary schools continue to allocate more experienced teachers in charge of students who are doing well due to external pressures such as league tables, instead of students who struggle academically, the primary education system is likely to further perpetuate educational inequalities with implications for social mobility.

No association between attainment grouping in primary school and externalising psychological symptoms during childhood and adolescence was found, regardless of initial verbal skills or socioeconomic circumstances. This finding contradicts the idea that differentiating learning environments for high- and low-attaining students in primary school reduces misbehaviour by reducing the possibility of distraction from both ends of the attainment distribution (House of Commons., 2011). Finally, attainment grouping in primary school was not associated with children's internalising symptoms during childhood and adolescence, regardless of initial skills or socio-economic circumstances. Limited support was then provided to recently raised concerns from primary school teachers in English primary schools that early attainment grouping or attainment grouping later in primary school could have a potentially negative effects upon some children's mental health (Bradbury and Roberts-Holmes, 2017).

Contents

List of Tables

Chapter 2

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix 3

Appendix 4

List of Figures

Chapter 2

Chapter 4

Chapter 5

Chapter 6

Equations

List of abbreviations

95% CI 95% Confidence Interval

BAS British Ability Scales

CAPI Computer-Assisted Personal Interviewing programme (CAPI)

CLS Centre for Longitudinal Studies

DfE Department for Education

EAL English as an Additional Language

GDP Gross Domestic Product

GSCE General Certificate of Secondary Education

GWAS Genetic Wide Association Studies

KS Key Stage

MAR Missing at Random

MCAR Missing Completely at Random

MCE Monte Carlo Error

MCS Millennium Cohort Study

MICE Multiple Imputation Chained Equations

MNAR Missing not at Random

NPD National Pupil Database

OECD Organisation for Economic Cooperation and Development

OR Odds Ratio

RRR Relative Risk Ratio

SATs Standardised Assessment Tests

SDQ Strengths and Difficulties Questionnaire

S.E Standard Error

SEN Special Education Needs

UK United Kingdom

US United States

Chapter 1: Introduction

1.1: My PhD journey

My PhD journey started long before I made my application to UCL Research Department for Epidemiology and Public Health and the ESRC. It started in my mind, as I gave some thought to why I wanted to pursue a PhD. I was curious about the knowledge I gained from the MSc Health and Society: Social Epidemiology at UCL in 2013/14 and I wanted to know 'how' and 'why' the social context could contribute to human development. Initially, I was interested in whether and how the neighbourhood environment influences children's cognitive and psychological development in the UK. However, a similar project had been completed by Dr. Anja Heilmann, who later became one of my PhD supervisors. My focus then shifted to the role of school environment on child development. I successfully applied to the ESRC and I was granted a three-year studentship to investigate how the quality and the composition of UK primary schools could contribute to children's cognitive and psychological development. Since school quality and composition data were not available in the MCS, I proposed to link data from Ofsted reports and data on school characteristics from the National Pupil Database (NPD).

I started my PhD in October 2015 and I successfully applied to the UK Data Service for secure access to already linked individual level NPD data to the MCS in April 2016. These linked datasets include measures of individual level attainment, absence and exclusion, and individual level demographics. Since there were on average 2 MCS participants in each participating school, it was difficult to aggregate these individual level data at the school level. Therefore, in May 2016 I applied to the Department for Education (DfE) to link school level NPD data to the MCS. Such data linkage would have allowed to investigate whether school composition, such as the proportion of pupils with Free School Meals or the school average attainment, were associated with MCS participants' cognitive and psychological outcomes. In December 2016, I successfully received training on accessing secure data through the UK Data Service Secure Lab. In January 2017, I submitted my upgrade report and I successfully passed my upgrade examination in February 2017, assessed by the UCL Institute of Education Emerita Professor Heather Joshi. In March 2017, aggregate school level data from the NPD and Ofsted reports data were sent to the CLS to be linked. However, technical, and legal issues delayed the data linkage to the MCS for several months. For this reason, in September 2017, I decided along with my PhD supervisors to change to my current PhD topic.

Since I was entering my third year of my PhD with only one year of funding left, I requested an extension of funding from the ESRC UBEL Doctoral Training Partnership. Fortunately, I was granted an extra year of funding until September 2019 to complete my PhD. This gave me the opportunity to complete the analysis for my project, write up my thesis and gain valuable academic experience by speaking at two international conferences: The Society for Life Course and Longitudinal Studies in Milan (July 2018) and the British Educational Research Association conference in Manchester (September 2019).

1.2: Social context of the research: Children's psychological symptoms and verbal skills in the UK.

In the UK the Office for National Statistics Child and Adolescent Mental Health Survey in 2017 estimated that 11.2% of children and young people aged between 5-15 years had a mental health disorder compared to 9.7% and 10.1% in 1999 and 2004 respectively, with the most common diagnoses being depression and anxiety, and conduct and hyperkinetic disorders (Sadler et al., 2018). 50% of mental health problems are established by age 14 and 75% by age 24 (Kessler et al., 2005). It is now well established from a variety of studies, that the presence of mental health illness can restrict someone's life chances regarding their future educational opportunities and work prospects, their physical health, their life expectancy, and their likelihood of committing a crime (Hemingway and Marmot, 1999, Nicholson et al., 2006, Rothon et al., 2009, Sayal et al., 2015, Hughes et al., 2020, Messer et al., 2004, Clark et al., 2017, Egan et al., 2016, Hope et al., 2019, Chang et al., 2011). Taking into account the individual and social costs, the Centre for Economic Performance's Mental Health Policy Group (2012) estimated that mental illness reduces the Gross Domestic Product (GDP) by £52 billion a year. A recent Organisation for Economic Cooperation and Development (OECD) report estimated that mental illness is costing the UK more than £94 billion a year or about 4.1% of the UK GDP, counting treatment, social support costs and the losses to the UK economy from people who cannot work (OECD, 2018b). Nearly half of the UK costs (£42 billion) are indirect costs related to higher unemployment rates and subsequent loss of productivity.

Contrary to mental health, the social determinants of verbal development have attracted very little attention from the epidemiological community in the UK. Verbal skills are a marker of cognitive performance. They refer to the extent to which a person uses words, sentences, texts, and grammatical rules to communicate. Moreover, they refer to the extent to which a person can understand meanings based on used words or complete sentences with words omitted dependent on the appropriate word context, and have a critical view towards written speech.(Newman and Newman, 2015, Wells, 2009).

According to the Bercow review (2008), 7% (nearly 40,000 children in 2007) of five year olds entering primary school in England had speech, language and communication difficulties, with this percentage declining with increasing age (Meschi, 2010). It is estimated that language, speech and communication difficulties affect approximately 10% of children and young people at some point in their lives (Norbury et al., 2016). However, these difficulties are not equally distributed among children and adolescents. A strong social gradient exists with language difficulties being more prevalent in children and young people from socio-economically disadvantaged populations compared to children and young people from more affluent populations (Bercow, 2008, Locke et al., 2002, Law et al., 2011).

Language, and especially the ability to communicate is an essential life skill for children and young people in the 21st century (Bercow, 2008). Social interaction is facilitated by communication. With effective verbal and communication skills, children can engage and thrive in life. Without them life could be a struggle for them. They will face difficulties in learning, achieving, making friends, and interacting with the world around them. A number of studies have shown an association between early life language problems and social difficulties, bullying or social exclusion by other children and adolescents without language problems (Lindsay, 2008, Botting, 2000, Markham et al., 2012, van den Bedem et al., 2018), with recent studies demonstrating a continuity between language difficulties in first years of life and adult social problems. One of these studies by Schoon and colleagues (2010) emphasized the importance of language difficulties during childhood and their negative association on later psychosocial adjustment in adulthood. Literature has also shown that young people with persistent language difficulties , who have had little support, have a lower exam pass-rate and fewer go on to further or higher education (Clegg et al., 2005, Conti-Ramsden et al., 2009, Department for Education., 2017), leading to restricted employment opportunities reflected in routine or semi-routine work (Feinstein and Bynner, 2004, Law et al., 2009, Conti-Ramsden et al., 2018).

The evidence presented in this section suggest a pertinent role for early verbal skills and psychological wellbeing for longer term social and health outcomes in adulthood with economic ramifications for the individual as well as the society. Therefore, a greater understanding of the determinants of verbal skills and psychological wellbeing in childhood and adolescence is crucial to averting an increasing social and economic burden to the UK population.

1.3: Rationale of the research

Existing evidence on the association between verbal skills as well as psychological symptoms in the early years of life and health outcomes in adulthood highlight the significance of environmental exposures during sensitive periods in childhood and the possible accumulation of risk over the life course. Therefore, studying environmental conditions in the early years of life that may influence the development of verbal skills and psychological symptoms over time is of great importance.

Environmental influences which are responsible for fostering nurturing conditions for children range from more proximal factors, such as the family environment to broader socioeconomic context shaped by governments, international agencies and civil society (Irwin et al., 2007). One context which has not gained much research interest in social epidemiological circles until recently is the school environment and its practices. To date, emphasis has been put on the role of secondary schools on cognitive and non-cognitive outcomes of adolescents (Kidger et al., 2012, Leckie, 2009, Gershoff et al., 2009), with little known about primary schools, their practices and their effect on child development in the western world, including the UK (Blanden et al., 2015, Education Endowment Foundation, 2018a).

The idea for this project was born following claims in 2016 by the former UK Prime Minister Theresa May, of a *'great meritocracy', 'where children have a fair chance in life, the chance to go as far as their talents will take them'* (May, 2016). These claims were made in an attempt to gain support from the UK electorate for the policy of opening new grammar schools that could cater differently for children according to their *''different talents, different interests and different dreams''* (May, 2016). The logic behind opening new grammar schools was based on the idea of meritocracy according to which secondary schools should allow every child to flourish and that all children have an equal chance to succeed.

International comparisons suggest that countries with lower ages of selection tend to show greater social disparities (Green, 2003 , Hanushek and Wossmann, 2006, Green et al., 2006) and reduced intergenerational mobility (Brunello and Checchi, 2006, Maurin and McNally, 2007). If attainment grouping exacerbates inequality during schooling years, earlier exposure to it may be more damaging to children's verbal and psychological outcomes.

Considering the above, this PhD aims to explore whether 'ability'[1] grouping in primary school is associated with better verbal and psychological outcomes in childhood and adolescence in the UK. Ability grouping hereafter referred to as attainment grouping. The role of attainment grouping on disparities in children's verbal and psychological symptoms according to family income and initial verbal skills will be also explored. The novelty of this PhD lies in being one of the first attempts to explore longitudinal associations between attainment grouping in primary school, family income, initial verbal skills and the development of children's verbal skills and psychological symptoms in a UK contemporary context.

1.4: Thesis outline

The overall structure of this thesis takes the form of eight chapters. Chapter 2 reviews the relevant literature, beginning with laying out the theoretical dimensions of this field of research, and looks at how attainment grouping has evolved over time in the UK. It also includes a review and a critique of the findings of past research on the relationship between attainment grouping, achievement tests and academic-related non-cognitive outcomes among primary and secondary school pupils in the UK. Chapter 3 begins with laying out the main contributions of this PhD and goes on to describe he aim, objectives and relating hypotheses of this PhD. Chapter 4 describes the data sources as well as the specific methods all statistical analyses of this PhD are based on. Chapter 5

[1] Although these practices are sometimes described as 'ability' grouping, the author of this PhD takes the position that 'ability' is not ascribed and fixed: rather it is malleable reflecting a range of societal factors that impact child outcomes. Hence 'ability' grouping is referred to as attainment grouping in this PhD project.

focuses on the association between the transition to attainment grouping between 7 and 11 years and the change in MCS participants' verbal skills over the same period. Chapter 6 shows findings on the associations between the transition to attainment grouping between 7 and 11 years and the change in parent- and teacher reported externalising and internalising symptoms between 7 and 11 years. Chapter 7 explores the long-term association between attainment grouping at age 7 and trajectories of parent-reported externalising symptoms from 7 to 14 years. This chapter aims at contributing to a better understanding on whether the trajectories of MCS participants who were grouped by attainment at age 7 changed considerably during the transition from childhood to adolescence. Studying the early adolescent period has the potential to convey more about the long-term effects of attainment grouping in primary school.

The final chapter (chapter 8) provides an overall discussion of the findings in relation to objectives and relating hypotheses of this PhD as well as their implications for future research and educational policy.

Chapter 2: Literature Review

2.1: Introduction

Throughout the world, national governments across the political spectrum place a great importance on their education systems. They embrace the idea that in a globalised technology-dominated world, a better educated and flexible work force is necessary (Brown et al., 2001). At the same time, in the light of climate emergency and the urgency to use the natural resources of the planet wisely, there are voices echoing the need for better education and the inclusion of environmental issues within the national curricula (Bangay and Blum, 2010, Fahey, 2012). Thus, it seems that both national governments and activists regard education as a powerful tool for changing attitudes of the future generations.

However, pupils may not only be affected by what they are taught in school, but also by the way learning is organised within school. This PhD will therefore attempt to explore whether the way learning is organised in UK primary schools is associated with child development. This PhD thesis comes at a time when the number of UK primary schools adopting attainment grouping has increased substantially (Hallam and Parsons, 2013a, Hallam and Parsons, 2013b). It is believed that this increase is driven by the salience of Standardised Assessment Tests (SATs) scores in defining school effectiveness and the attainment grouping may increase student's academic attainment. Important is also the notion that in a competitive world, a competitive education system can best accommodate the needs of the cleverest pupils by keeping them together and giving them a different learning experience (Crozier et al., 2011, Vincent and Ball, 2007).

In the UK, streaming and setting are two of the most common forms of attainment grouping. They can be practiced alone or in tandem (Hallam and Parsons, 2013a, Hallam and Parsons, 2013b). They are-based on the theory that every pupil can learn and make progress according to their ability, along with a group of similarly able pupils (Education Endowment Foundation, 2018a). This theory has been widely accepted by successive

UK governments, educators and parents (Francis et al., 2017a). However, justification for this theory can be undertaken through empirical examination of the role of attainment grouping on pupils' development of verbal skills and psychological symptoms.

The first section of this chapter defines attainment grouping and gives a brief outline on two hypotheses (the divergent and the polarisation hypotheses) that may explain the potential impact of attainment grouping on students' outcomes. This PhD focuses on the potential impact of attainment grouping on children's developmental outcomes. Focus is placed upon children's cognitive skills and psychological well-being. For this reason, the life course theory is chosen as the overarching framework of this PhD to describe the effects of multiple environmental risks on child and adolescent developmental outcomes. Then, it discusses the ecological models of human development theory described by Urie Bronfenbrenner. The ecological models of human development theory describe human development as an interaction between different contexts such as the interaction between the family and the school environment. Further, literature is reviewed with an exclusive focus on the UK context, where few published studies exist on the association between attainment grouping, cognitive and psychological outcomes. The literature review begins with an overview of the historical and contemporary socio-political context of within-school selection in the UK education system. It further presents existing evidence about the role of attainment grouping practices at primary and secondary school level on children's and adolescents' cognitive, social, and emotional development in the UK. Finally, the limitations of previous UK research are discussed.

2.2: What is attainment grouping?

In the UK most contemporary debates about selection in education are focused around fee-paying versus state-maintained education or between selective (grammar schools) and comprehensive secondary education. However, both fee-paying and grammar schools are only a small part of the education system in the UK (Andrews et al., 2016, Department for Education., 2019a). The UK leads the world in the proportion of students separated within their schools into different classes based upon selection on academic attainment (streaming or setting) (OECD, 2013).

Several types of within-school selection have been mentioned already. For clarity, Table 2.1 provides a summary of the main types of attainment grouping used within UK primary schools.

Table 2.1: Types of attainment grouping

Streaming	**Pupils are placed in classes based on a test of their general cognitive ability. They remain in their streamed class for most subjects.**
Setting	Pupils are grouped according to their attainment in a subject (e.g. English, mathematics, science). Setting may be imposed across a whole year group, across timetable halves, within a stream or across mixed age classes. Sets may be hierarchically ordered or there may be parallel sets.
Mixed attainment	All students are taught as one group in the same classroom regardless of their cognitive ability.
Within-class attainment grouping	Pupils are grouped within the class because of their cognitive ability. Within-class groups may be regrouped for different subjects

Streaming is the most rigid form of attainment grouping. Pupils are placed into a class based on a measure of their overall cognitive ability and they are taught in that class for most subjects. Streaming assumes that students have a fixed level of general intelligence, which predicts their performance across all subjects, and can be measured by objective tests (Ireson and Hallam, 2001). Streaming may also be adopted in order to homogenise groups of students according to their general intelligence, thus making it easier for teachers to match their teaching to pupils' levels of academic ability (Hallam and Parsons, 2013b).

Setting is a more flexible type of attainment grouping where pupils are grouped according to their academic attainment in a particular curriculum subject (Ireson and Hallam, 2001).Setting is based upon the assumption of differential ability in a particular curriculum subject/s which allows for higher academic attainment for some pupils. Schools may use setting for some or all academic subjects. Similar to streaming, setting aims to reduce the spread of ability in a particular curriculum subject within class and enable teachers to match their teaching to pupils' needs (Hallam and Parsons, 2013b).

Setting and streaming are appealing organisational practices to UK primary schools since they have few associated financial costs. However, if streaming or setting results in larger number of classes or requires additional resources for different groups, then additional expenditure may be needed (Education Endowment Foundation, 2018a).

Within-class ability grouping is the most popular and widely adopted type of grouping within UK primary schools, especially in early primary school years (Education Endowment Foundation, 2018b). It differs from streaming or setting in the sense that groups can be organised within class according to pupils' general cognitive ability or their academic attainment on a particular subject, or it may be based on working relationships between pupils (Bradbury and Roberts-Holmes, 2017).

Mixed-attainment grouping refers to the random assignment of students in a class to balance the mix of abilities in each class by using available achievement test scores. Other considerations may involve a balance of boys and girls in a class, or a similar balance of ethnic minority pupils (Sukhnandan and Lee, 1998).

2.3: The effects of attainment grouping on pupils

There are two main hypotheses on how attainment grouping practices may affect cognitive and psychological outcomes in childhood and adolescence. The first is the divergent hypothesis, and the second is the differentiation-polarization hypothesis.

2.3.1: The divergent hypothesis

The divergence hypothesis suggests that homogeneous attainment group teaching is associated with a divergence in pupils' attainment outcomes and an increase in differences between lower and higher attaining pupils, while heterogeneous mixed-attainment teaching is associated with either a smaller increase, no increase or a reduction in attainment differences. In his study Kerckhoff (1986) compared reading and mathematical test performance both between different types of schools (secondary modern, comprehensive, fee-paying and grammar) and within schools, comparing those who were grouped by attainment and those who did not. The study's findings supported the divergence hypothesis and showed that differences in attainment increased for pupils taught in attainment groups. The study attributed the increase reading and mathematics attainment differences to the negative effect of attainment grouping teaching on the attainment of lower attaining pupils and a small positive effect on the attainment of higher attaining pupils. In another study Lacey (1970), (Lacey, 1974) conducted a two part study in a school. At first, the school was teaching its pupils in streams and then (in the follow up study) adopted mixed attainment teaching. The study highlighted that the adoption of mixed attainment teaching led to improvements in attainment for the lower attaining pupils without having a negative effect on the attainment of higher attaining pupils. A study by Linchevski and Kutscher (1998) compared the academic progress of grade 9 (age 13/14 years) pupils in mixed attainment and attainment grouping classes for mathematics. A divergence in mathematics attainment was observed with lower attaining pupils facing the greatest losses (a slower progress in mathematics than higher attaining pupils). Other studies exploring progression from KS2 to KS3 and KS3 to GCSE

(Venkatakrishnan and Wiliam, 2003b, Wiliam and Bartholomew, 2004, Ireson et al., 2005) provide further support for the divergence hypothesis with pupils in lower attainment groups making less progress. The divergence hypothesis is important as it suggests that a divergence in pupils' attainment within classes grouped by attainment happens because of a slower progress for lower attaining pupils rather than enhancing the achievement of higher attaining pupils. Alternatively, the reduced divergence in pupils' attainment in mixed attainment groups is likely to be attributed to the improved/faster progress of lower attaining pupils than depressing the progression of high attainment pupils.

2.3.2: The polarisation hypothesis

The polarisation hypothesis suggests that attainment grouping may result in a polarisation of pupils' behaviour and attitudes because of attainment grouping, both within and between groups and towards school (Berends, 1995, Van Houtte, 2006, Muller and Hoffman, 2016, Gamoran and Berends, 1987, Van Houtte, 2016).

Polarisation is apparent in all types of attainment grouping (within-class groups, streams, or sets). Stigmatisation, name-calling and teasing from other pupils are common in schools adopting attainment grouping (Hallam et al., 2004b). For instance, high attaining students may be called “brainies”, “goody-goodies” or “whiz kids”, while low-attaining students may be labelled as “thick”, “slow”, “difficult”, “dumb”, “crap” and stupid”. According to a six-year old student in a London primary school (Reay, 2017):

> *“They [the Lions] think they are better than us. They think they are good at every single thing and the second group, Tigers, there are some people that think they are good and more important than us. And one of the boys in Giraffes he was horrible to me and he said “get lost slow tortoise” but my group are Monkeys and we are only second to bottom”*

Further, attainment grouping seems to be associated with stereotyped descriptions by teachers. In England, low-income pupils seem to be underestimated by their teachers, along with pupils with any SEN diagnosis, ethnic minority pupils, pupils speaking languages in addition to English and boys (reading)/girls (maths) (Campbell, 2015). In addition, students with these characteristics are more likely to be allocated to bottom sets where teacher practices and expectations vary compared to higher streams and sets where higher-income, White, non-SEN and English speaking students are more likely to be allocated (Hallam and Parsons, 2013a, Hallam and Parsons, 2013b, Hartas, 2018). As a result curriculum differentiation is polarised between streams and sets, with, for instance, students allocated to top sets for mathematics being ascribed qualities as "mini-mathematicians", not as a result of their individual qualities, but simply due to their location in a top set (Boaler et al., 2000). In her study, Boaler (1997b) found that teachers change their normal practices when they are assigned top-set classes, appearing to believe that being a 'top-set' student entails a qualitative and meaningful difference from other students, rather than simply being in the highest-attaining range of students in the school. Teachers in the study seemed to believe that top-set children did not need detailed help, time to think, or the space to make mistakes. Rather, they could be taught quickly and procedurally because they are clever enough to draw their own meaning from the procedures, they are given. As it was put by a girl in a top set for mathematics;

> *"Cos you learned a lot more [in mixed-ability groups] and you could recap everything which you didn't understand and spend more time on it, but now [in set 1-top set] you've just got to try and whizz and do your best."*

These stereotypes are not without consequences as they seem to be associated with a negative impact on students' enjoyment for mathematics. In the same study, students in top sets (43%) reported not enjoying maths lessons, compared to an average of 36% of students in other sets, and 32% of students in mixed attainment classes (Boaler, 1997b).

It seems that teachers are responding to their students' academic needs by grouping them in homogenous classes accordingly. As discussed above this may have negative effects not only on pupils in higher groups, but also for students in lower attainment groups who may consider schoolwork less demanding (Boaler et al., 2000, Ireson and Hallam, 2001). Nevertheless, one important function of reducing the level of challenge in lower attainment groups may be that it simultaneously reduces the risk of failure for both teacher and pupils (McManus, 2010). As Schwartz (1981) put it;

“High rank pupils who are likely to succeed and contribute positively to the teachers' professional image are perceived as ideal pupils whose specific educational needs the teacher is capable of meeting. By contrast teachers tend to distance themselves from lower ranked pupils viewing them more as an unreachable group than as a series of individuals with distinctive educational problems. Rather than risk professional failure with pupils whom they fear they will be unable to motivate teachers often make fewer demands on low-rank pupils and apply less exacting standards to their own performance with them.”

Attainment grouping also seems to play an important role in peer relationships. There is some evidence that membership in an attainment group has a positive effect on the probability that a student will choose a peer as a best friend, however the formation of cliques between attainment groups may increase over time (Hallinan and Sorensen, 1985). This can lead to the development of pro- and anti-school attitudes and behaviours (Berends, 1995). Of interest is the emergence of anti-school subcultures as students in low attainment groups, stereotyped and alienated from high attaining peers, may seek non-academic routes to maintaining their self-esteem; according to a bottom set student in an English comprehensive school (Reay, 2017);

“The behaviour, it gets worse in the bottom set when, like teachers don’t pay attention to you. And they pay attention to, like, the higher ability students and, like you get bored because there is nothing for you to do if you don’t understand the work”.

Students in lower attainment groups may feel that, they will have limited future job prospects later in their lives. Such feelings can lead to an individual sense of futility since students start to believe that they cannot succeed in school anymore (Brookover et al., 1978). In turn, a high sense of futility may be related to school misconduct (Van Houtte and Stevens, 2008).

Finally, there is some evidence of polarisation in peer relationships within groups. Students in lower attainment groups might be more hostile towards each other. The words of a primary school student are characteristic (Hallam et al., 2004b);

"I'd put the brainy ones not near me but where I could still see them and the people who don't act their age and they do know a lot but they don't get on with it….they'd be near me so if they don't get on with it I could tell them off"

According to Schwartz (1981) low attaining students *"turn their negativity about their status onto each other"* and says that:

"[For low ability students] caught in a dilemma by their academic label, it becomes more important to compete with and differentiate oneself from like-ranked peers than to complete the task at hand. By downgrading others' efforts and intelligence, one can set oneself apart from classmates and ensure that they do not succeed where one might fail."

2.4: Dimensions of human development

Human development refers to physical, cognitive and psychological changes that children undergo from their birth to their adolescence (Trawick-Smith, 2008, Jensen Arnett, 2012). This PhD will pay attention to cognitive and psychological outcomes during childhood and adolescence.

2.4.1: Cognitive skills

Cognitive skills have long been considered the stepping stone of success in adult life through their contribution towards optimal labour market outcomes (Joshi, 2014, Currie, 2001, Carneiro and Heckman, 2003). Cognitive skills are multifaceted and encompass a series of skills ranging from reasoning, memory, problem solving, abstract thinking, idea comprehension, and learning from experience (Gottfredson, 1997).

In the past, these skills were thought to be substantially heritable, approximately 50% based on the study of twins (Plomin and Spinath, 2002, Krapohl et al., 2014, Scourfield et al., 1999, Toga and Thompson, 2005, Deary et al., 2009), but there is growing evidence contesting the dominant purely genetic component of cognitive skills (Savage et al., 2018, Hill et al., 2019). For instance a recent Genome Wide Association Study (GWAS) using cognitive and genetic data from the CHARGE and COGENT consortia, and the UK Biobank study from 300,486 participants found that generated polygenic scores[2] accounted for only 4.3% of variance in general cognitive function (Davies et al., 2018). Another study using actual data on genes and cognitive scores from the ALSPAC cohort study found three candidate genes to be associated with reading scores, which accounted for only 2% of the variance in children's scores (Jerrim et al., 2015). Given that the genetic component of cognitive skills has been found to be less than it was previously thought, that allows for a greater influence of environmental factors. It is believed that cognitive skills can be cultivated by appropriate training, incentives and challenges (Dweck, 2007).

Cognitive skills or intelligence can be distinguished to fluid or crystallized intelligence. Fluid intelligence is the ability used in inductive and deductive reasoning, particularly with novel material. Inductive reasoning moves from specific instances into a generalised conclusion e.g every quiz has been easy, so this test will be easy. Deductive reasoning moves from generalized principles to specific e.g quadrilaterals have four sides, so a square is a quadrilateral. Fluid intelligence can be contrasted with general crystallized ability which reflects schooling and acculturated learning, and the two abilities have different developmental trajectories. Crystalized intelligence involves knowledge which comes from learning and previous experiences (Barbey, 2018). Situations that require crystallized intelligence include reading comprehension and vocabulary.

2.4.1.1: What the cognitive tests are testing?

Cognitive tests measure similar cognitive abilities and processes, but typically for slightly different purposes (Anastasi, 1982). Achievement or performance tests require previous study to get a better score, while aptitude tests demand no previous study. Aptitude tests assess only the current level of cognitive ability. Another difference is that

[2] A polygenic score is a number based on variation in multiple genetic loci and their associated weights. A polygenic score serves as the best prediction for the trait that can be made when considering variation in multiple genetic variants.

aptitude testing is designed to measure in which fields there is ability for growth, whereas achievement tests assess the knowledge already obtained in a field. Aptitude tests are therefore useful in helping to make informed decisions on the future by diagnosing someone's potential, while an achievement test helps to understand the progress made so far by a student. Conceptually, achievement tests are usually limited to assess knowledge in a certain field, whereas aptitude tests are designed to make a general evaluation.

2.4.2: Psychological development

In 2017, 11,2% of children and adolescents aged 5-15 years-old could be suffering from a clinically diagnosed mental disorder in the UK, with the most common diagnoses being depression and anxiety, and conduct and hyperkinetic disorders (Sadler et al., 2018).

A well-known distinction in the field of child psychology and human development is the distinction between externalising and internalising symptoms (Achenbach and Edelbrock, 1978). The construct of externalising problem behaviour refers to a grouping of symptoms that are manifested in children's outward behaviour such as attention deficit or/and hyperactivity and conduct symptoms. The construct of internalising problem behaviour reflects the grouping of emotional and anxiety symptoms.

Investigating externalising and internalising symptoms in children and adolescents is crucial since there is empirical evidence of continuity of psychological symptoms to adulthood Externalising symptoms have been found to be associated with a range of adverse outcomes such as leaving school earlier, teenage pregnancy, lower educational attainment, unemployment, financial difficulties and substance use (Esch et al., 2014, Liu, 2004, Moffitt and Team, 2002, Colman et al., 2009). Internalising symptoms in childhood or adolescence have also been found to be associated with poor academic achievement and adult depression (Jakobsen et al., 2012, Riglin et al., 2014, Okano et al., 2019, Olivier et al., 2018, Moffitt et al., 2007, Colman et al., 2007).

2.5: Theories relating to Human Development

The overarching framework of this PhD is the life course theory (Elder, 1998). The theory places child development within a social, cultural, and historical context. The literature shows that childhood and adolescence are important life stages in the development of cognitive skills and psychological symptoms. Work in the area of human development, has also identified that family and neighbourhood contexts, as well as relationships with peers are critical for development during these life stages (Heilmann et al., 2013, Flouri et al., 2019a, Flouri et al., 2017, Flouri et al., 2019c, Mueller et al., 2019, Ahn et al., 2018, Flouri and Midouhas, 2017, Flouri et al., 2019b, Flouri et al., 2015a, Flouri et al., 2015b, Midouhas et al., 2014, Flouri et al., 2014, Harding et al., 2015, Del Bono et al., 2016, Kelly et al., 2013c, Kelly et al., 2013a, Kelly et al., 2011, Cooper, 2017). However, the role of school environment and its practices as antecedents of development remain less clear (Heilmann, 2013, Papachristou et al., Mihoudas, 2017, Flouri and Midouhas, 2016, Kidger et al., 2012). In ecological systems theory (Bronfenbrenner, 1977), at any given life stage, individuals interact with their immediate changing contexts. These interactions include social and material, as well as influences from their social class, cultural and societal norms. How primary schools organise the learning of their students may be particularly important for their verbal development and externalising and internalising symptoms over time. Therefore, both the life course and the ecological systems theories inform the analyses of this PhD thesis.

2.5.1: The life course theory

A theoretical perspective which must be noted for influencing the conceptual framework of this research is the life course theory. The life course can be characterised as a multi-disciplinary framework within which all explanatory models (behavioural, psycho-social and materialistic) can be encompassed (Bartley, 2017). The life course approach is extensively used in population health research since it provides an explanation of the observed social gradients in health outcomes (Marmot and Wilkinson, 2006). Social gradients are the result of complex combinations of peoples' circumstances taking place over their lifetime (Ben-Shlomo and Kuh, 2002).

A life course approach in epidemiology *"studies the physical and social hazards during gestation, childhood, adolescence, young adulthood and midlife that could influence chronic disease risks and health outcomes later in life"* (Kuh et al., 2003). A life course framework can help us understand the biological, behavioural, and psychosocial factors throughout one's life, which can independently, cumulatively, and interactively influence health. Such a consideration can also be useful in identifying possible risk and protective factors affecting health and social outcomes (Ben-Shlomo and Kuh, 2002). From a life course perspective, a developing person's past experiences relate to stable and long-term changes in bio-developmental stages (Hertzman, 2012). The social is literally embodied. Children in the UK start their lives from very different positions. A combination of their parents' social position at birth, as well as the structured nature of social processes taking place at home and the school environment over time contributes to accumulation of future advantage or disadvantage (accumulation of risk model) in terms of their cognitive, social and emotional development (Waldfogel and Washbrook, 2010, Lareau, 2011, Ben-Shlomo and Kuh, 2002, Hertzman and Boyce, 2010). An auspicious start in life, as a mark of social distinction predisposes to better cognitive outcomes and lower psychological symptoms in the future. It may translate to future advancement and social advantage, while early life disadvantage predisposes to less favourable cognitive and non-cognitive outcomes, thus to future deprivation and social disadvantage.

A key element then in life course epidemiology is the temporal ordering of exposures and their inter-relationships (Kuh et al., 2003, Ben-Shlomo and Kuh, 2002). Time is an important concept in life course research, both in terms of lifetime (chronological age of individuals) and historical time at the population level (membership of a birth cohort). The effect of an exposure on a health or a social outcome may be dependent on the duration or timing of exposure. A critical period model is so defined because the age of the individual at the time of exposure is assumed to influence the long-term outcome. In other words, the critical period model reflects qualitatively different exposure-time interactions (Kuh et al., 2003) . Evidence suggests that there are critical periods of growth and development, not only in utero as supported by the ‘fetal origins hypothesis’ and early infancy (Barker, 1990), but also during childhood and adolescence. In critical periods, environmental exposures do more damage to health, long-term health and social potential than they would at other times (Steinberg, 2005b). There is also evidence about sensitive developmental stages in childhood and adolescence where cognitive and non-cognitive skills, habits, coping strategies, attitudes, and values are more easily acquired than at later ages. For instance, a recent US study of almost 700,000 native and non-native English speakers found that the grammar learning ability for a second language acquisition is stronger during the first 17-18 years of life and then it starts declining (Hartshorne et al., 2018).

Further, a critical period model with later effect modifiers implies investigating plausible interactive effects between early and later life risk factors. In social epidemiology, it is argued that a risk or a protective factor modifies the association between an exposure and a health outcome when the causal effect of the exposure of interest differs across levels of the modifying factor. Investigation of modifying factors provides information about the nature of the causal process (Kuh et al., 2003). In the language of statistics, effect modification is referred to as interaction— “synergism”, if the modifying variable enhances the effect of the explanatory variable, or “antagonism” if it diminishes it. Interactions are common features of life course processes and it is suggested that they should be investigated where plausible biological, behavioral or social hypotheses exist (Kuh et al., 2003).

Recognising the importance of accumulation, timing of exposure to attainment grouping, and interactions with individual and family characteristics such as family income and initial verbal skills are inquiries central to the current PhD.

2.5.2: Ecological systems theory

This PhD thesis also draws upon the theoretical underpinnings of the ecological model of human development described by Bronfenbrenner. According to Bronfenbrenner, in understanding human development, the entire ecological system in which growth occurs should be taken into consideration (Bronfenbrenner, 1977). According to ecological systems theory, human development is driven by the interaction of proximal processes, for instance the immediate environment in which the developing individual lives (family) and learns (school) (Bronfenbrenner, 1979, Brandt and Bronfenbrenner, 1979). Applying the influences of the immediate environment of the developing person described by Bronfenbrenner, this PhD thesis focuses on the interaction between the developing child, the family, and the school environment over time. The way UK primary schools organise the learning of their students (attainment grouping) may play a role in the development of children's verbal skills and psychological symptoms. However, it is possible that the influence of attainment grouping in primary school on children's development may be different according to family socioeconomic circumstances and initial verbal skills, an indicator of child's school readiness. These interactions taking place in the child's immediate environment may further be influenced by interactions in wider environmental contexts. These interactions range from social relationships in the area level to political, social and economic processes of a given geographical area (Bronfenbrenner et al., 1984, Bronfenbrenner and Evans, 2000). Therefore, perceiving, the ecological environment in which human development takes place as an interdependent "set of nested structures", moving from the innermost level to the outside (see Figure 2.1) is important in understanding whether attainment grouping plays an important role in the development children's verbal skills and psychological symptoms in the UK.

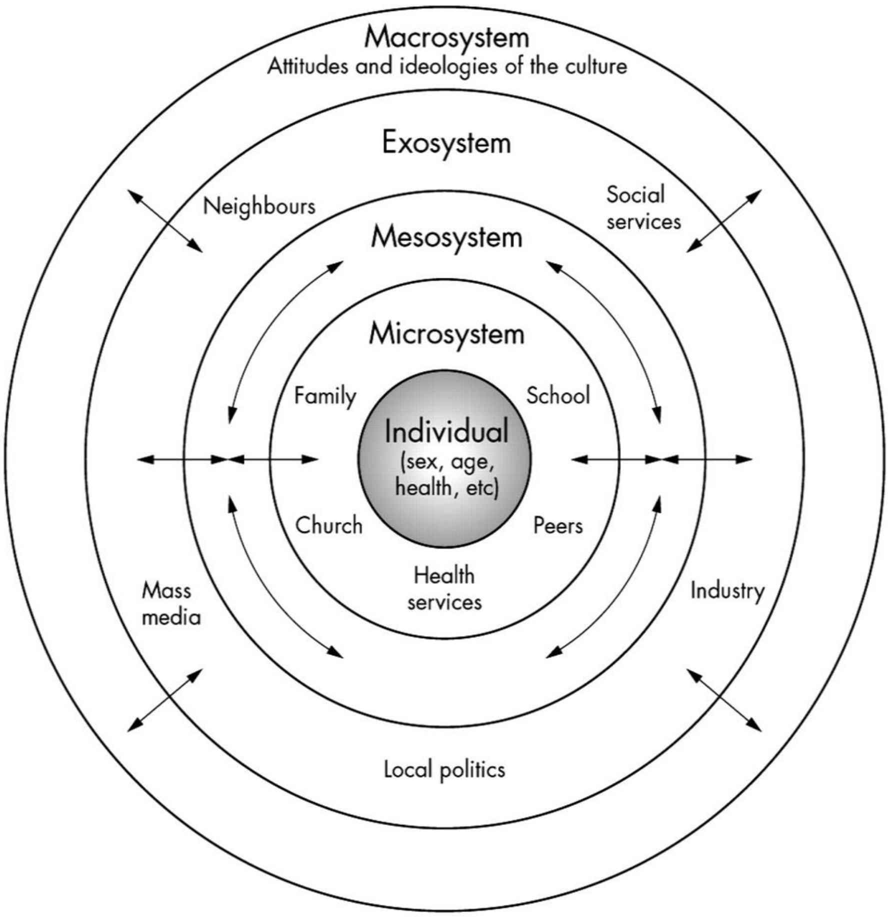

Figure 2. 1 Bronfenbrenner's ecological theory of human development

2.6: The role of age and gender on human development

Individual-level factors such as age, gender and early cognitive ability are important determinants of cognitive and psychological outcomes over time. Evidence shows that cognitive skills change with age and instruction (Steinberg, 2005a, Carneiro and Heckman, 2003, Carlsson et al., 2015, Kautz et al., 2014) Externalising and internalising psychological symptoms also vary by age. Evidence from the MCS shows that externalising psychological symptoms are more frequent during the early years of life and they decrease with time (Flouri et al., 2014, Patalay and Fitzsimons, 2017). In contrast internalising psychological symptoms are less frequent during childhood (Flouri et al., 2014), but they increase considerably during adolescence (Patalay and Fitzsimons, 2017).

Gender differences in cognitive as well as psychological outcomes are evident through childhood and adolescence. To identify the extent of any gender differences, Moss and Washbrook used data from the MCS and compared the percentage of boys attaining below the expected standard in verbal development at age 5, 7 and 11 with the percentage of girls (Moss and Washbrook, 2016). At each assessment point, they found between 12% and 6 % points more boys than girls were attaining below standard level. The gender gap remained relatively constant across all social class groups. In comparison, the initial gender gap in maths attainment was much smaller and became insignificant from age 7 onwards (Patalay and Fitzsimons, 2017, Sadler et al., 2018). Further, boys exhibit consistently greater levels of externalising symptoms across childhood and adolescence compared to girls. In contrast, girls seem to exhibit greater levels of internalising symptoms during adolescence compared to boys (Patalay and Fitzsimons, 2017, Sadler et al., 2018).

2.7: Family environment and human development

By the time children start primary school in the UK (age 5 years), there are pronounced differences in their cognitive and psychological outcomes, highlighting the importance of family environment in the early years of life (Kelly et al., 2011, Bradbury et al., 2015).

There is ample evidence on the association between parenting and human development. Particularly, authoritative parenting (solid but fair) is associated with more positive outcomes in children and adolescents (Galambos et al., 2003, Flouri et al., 2015c). Cognitive stimulating parenting behaviours are also important for cognitive and psychological outcomes. Frequency of reading, free writing, telling a story, crafts, painting, drawing, or drama are associated with more optimal cognitive and psychological outcomes (Fancourt and Steptoe, 2019, Burgess et al., 2006, Mak and Fancourt, 2020). Regular bedtimes or family mealtimes are related to improved cognitive and psychological outcomes (Kelly et al., 2013b, Kelly et al., 2013a, Elgar et al., 2013, Fiese and Schwartz, 2008).

Psychosocial factors related to family environment are important. There is evidence that parental relationship and parental divorce are associated with children's cognitive and psychological development (Mariani et al., 2017, Strohschein, 2005, Jekielek, 1998, Kim, 2011). The relationship between children and their parents is also decisive (Creese, 2019, Moullin et al., 2014). A vast body of literature has also highlighted that parental mental health, especially maternal mental health, and parental physical health are associated with children's cognitive and psychological outcomes (Coope, 2013, Creese, 2019, Kiernan and Huerta, 2008, Kiernan and Mensah, 2009, Mensah and Kiernan, 2010, Patalay and Fitzsimons, 2016, Fitzsimons et al., 2017b).

Whilst parenting practices and family psychosocial factors may be important for children's development, socioeconomic circumstances may be important for shaping parenting practices, family relationships, and parental mental and physical health which in turn may affect cognitive and psychological outcomes (Kiernan and Huerta, 2008, Kiernan and Mensah, 2009, Dickerson and Popli, 2016, McLeod and Shanahan, 1993, Cooper and Stewart, 2013a, Cooper and Stewart, 2017, Cooper, 2017).

2.8: Within-school selection in the UK

The next section of this chapter will attempt to provide a brief account of the historical context of attainment grouping in the UK and will review the existing scientific literature on attainment grouping and children's development of verbal skills and psychological symptoms in the UK.

2.8.1: Historical context

In the UK the popularity of attainment grouping has fluctuated significantly over the past century (Ireson and Hallam, 2001, Boaler, 1997a). It can be argued that public support for educational selection can be divided into three different periods; the pre-second World War, the post second-World War economic expansion (known as the golden age of Capitalism) and the rise of neo-liberalism in the early 1980's.

British society has always been divided along social class lines (Goldthorpe, 2016). For centuries schools were educating, almost exclusively, children from professional and managerial social classes. In the 19th century a series of changes in society and the economy, such as the expansion of the right to vote, the increasing need for specialisation of the workforce in industry, as well as the intensification of commerce and international trade urged the UK governments to establish a state-maintained comprehensive education system (Gillard, 2009).

By the beginning of the 20th century, the elementary education system had been established, with most children attending schools up to the age of 12 years. After the age of 12, only a small proportion of children continued their education in grammar schools. Most of them were coming from affluent families (Andrews et al., 2016). Developments in psychological theories of learning and eugenics provided a justification for segregating students (Ireson and Hallam, 2001, Lowe, 1997, White, 2006). A greater emphasis on cognitive testing with an ever-increasing refinement of tests of intelligence, aptitudes and competency provided a strong argument for those who believed that their success was solely due to merit. This view was compatible with the rationale for a selective education system and the segregation of students through streaming within primary schools.

An additional justification for not offering a common curriculum to all children in a well-resourced state elementary education system was due to preparing the UK's future workforce. A workforce ready to meet the needs of the industrialised UK society of the first half of the 20th century. The adoption of attainment grouping then by UK schools was a stepping stone to preparing a large number of future unskilled labourers who could be part of the production system, and smaller numbers of professionals and managers (Ireson and Hallam, 2001, Willis, 1978).

By the end of the World War II significant changes to UK society and the economy came into force which led to expansion of secondary education for all children. This expansion was accompanied by the loss of support for grammar schools and the stratification (grouping) of students within schools by attainment. There were multiple reasons for this transformation of the UK education system, antithetical to its old principles. During the 1960's and 1970's mounting evidence pointed to the failure of selective secondary education to serve as a means of equality of opportunity in education. A smaller number of students from working class backgrounds were able to enter grammar schools. Besides, the notion of intelligence as highly inherited was questioned, with evidence highlighting an ever increasing importance of the environment (Gallagher, 1994). There was also increasing concern that tests for intelligence aiming to select children on the basis of ability were unreliable, hence the abolition of 11 plus exams (Ireson and Hallam, 2001, Gillard, 2009, Miller-Jones, 1989).

The content of the Plowden report published in 1967 was also revolutionary with its clear encouragement for the un-streaming of primary education (Blackstone, 1967). There was also a change in educational values towards a child centred-education rather than whole class teaching inspired by newly formulated theories of influential educationalists such as Piaget, Montessori and Froebel (Gillard, 2009). These theories dictated that children's learning was different from that of adults and required much more active exploration of the world and freedom to develop at their own pace (Kitson and Merry, 1997). For this reason, many streamed primary schools reorganised their students to mixed attainment classes until the late 1990's. Many primary schools adopted more flexible forms of attainment grouping such as setting during that period. By the 1990's less than 3% of schools were streaming their students (Lee and Croll, 1995).

2.8.2: The current context

After the 1980's, economic forces related to de-industrialisation and the rise of a globalised knowledge-based economy provided the ground for a resurgence of support for attainment grouping within UK primary and secondary schools as a means to raise academic standards (Ireson and Hallam, 2001). Decisive determinants for this resurgence were the introduction of the National Curriculum and the 1988 Education Reform Act. Competitive market principles were introduced into the education system which endorsed parental choice and the change of the funding system tied to student numbers. Parental choice and competition between schools to attract a great number of students became more pronounced within England (Machin et al., 2013). These changes increased the pressure on UK schools to adopt attainment grouping (Boaler, 1997a, Marks, 2016). Therefore, the creation of an educational market has also encouraged within school selection, mainly setting, as schools compete to attract and retain students from middle class families (Ball et al., 1994, Whitty and Power, 2001, Francis et al., 2017a).

Pressures to raise the level of students' attainment in the UK were intensified by the publication of international comparisons of performance in mathematics and science, such as Programme for International Student Assessment (PISA). Such comparative studies have demonstrated over consecutive years that British students' achievement , especially those from disadvantaged socio-economic backgrounds is lower than students from other countries (OECD, 2018a). The low performance of British students forced the New Labour government after 1997 to lead initiatives to improve numeracy and literacy with the introduction of the National Literacy and Numeracy Strategies (Ireson and Hallam, 2001). These changes were accompanied by further recommendations for schools to adopt setting by 'ability'. Characteristically the white paper *"Excellence in Schools"* suggested that *"setting could be beneficial in raising standards and that setting should be the norm in secondary schools. In some cases it is worth considering in primary schools"* (Department for Education and Employment., 1997).

Over the last two decades, survey data suggest a considerable increase in the incidence and the prevalence of attainment grouping in UK primary schools. A survey based on five Local Authorities, which was carried out in 1999 with 111 participant primary schools, found that all reception classes and most Year 2 and Year 5 classes were mixed-attainment classes and only a quarter adopted setting and none of them adopted streaming (Baines et al., 2003). A survey of almost 800 randomly selected primary schools in 2000 highlighted that setting was more common in mathematics than English or science with the incidence rising from 1% in reception classes to 14% in Year 6 (Hallam et al., 2003). Findings from 2007-2008, using data from the MCS, also suggest that the incidence of streaming and setting in English and mathematics might have increased considerably in the UK; 37% of MCS children at age 7 were taught in sets for English or Mathematics, while a significant proportion of MCS children were taught in streams (16.5%) (Hallam and Parsons, 2013a, Hallam and Parsons, 2013b). Interestingly, a recent 2017 survey on 1373 education professionals related to Early Years education (ages 3 to 7) in England revealed that attainment grouping is quite prominent even among three year-old children; with 81% of reception teachers responding that they used grouping for phonics, while 58% of Nursery teachers used grouping for phonics and 35% for maths (Bradbury and Roberts-Holmes, 2017).

2.9: Searching strategy; Attainment grouping and cognitive skills

No individual UK studies on attainment grouping aptitude tests were identified in the last two decades via searching the Psych Info, ERIC and Educational Abstracts databases or searching manually via Google scholar. The reviewed studies and the findings reported focus on the association between attainment grouping and achievement tests (performance tests as they are known) such as the Key Stage (KS) exams[3]. Both performance and aptitude tests describe two related forms of cognitive testing as it was mentioned in subsection 2.3.1.1.

2.10: Attainment grouping in UK primary school and achievement

As it usually happens in debates around selection in education, whether the focus is placed upon the impact of between-schools selection (either fee-paying vs. state-maintained schools or grammar schools vs. comprehensive secondary schools) or on

[3] A Key Stage is a stage of the state education system in England, Wales and Northern Ireland.

within-school selection effects, two main questions are raised (Parsons et al., 2017, Gorard and Siddiqui, 2018). The first is whether educational selection raises educational attainment for all children. The second is whether it exacerbates socioeconomic and educational inequalities.

Only three studies on attainment grouping in primary school and achievement testing in a UK contemporary context (over the last two decades) were identified. Whitburn (2001) studied the progress of 1200 students in mathematics (200 students taught in set classes for mathematics and 1000 in mixed-attainment classes) from Year 2 to Years 3 and 4 of primary schools in the London Borough of Barking and Dagenham and three other boroughs in England. He found that when the same mathematics materials were used to teach primary school pupils in set and non-set groups, students in non-set classes improved their maths test scores up to 7% compared to students in maths sets. Whitburn also found that low-attaining students in mixed-attainment classes benefited more than low-attaining students in set classes, while high attaining students progressed at similar pace in Mathematics in both set and mixed-attainment classes.

In case studies of 24 primary schools covering four geographic regions in England including urban, suburban and rural areas, Kutnick and colleagues (2006) compared the KS2 test scores of Year 6 students who attended primary schools which were using setting (12 schools) and those which were not (12 schools). They concluded that setting rarely linked to higher KS2 scores when comparing local authority and national averages in English and Mathematics. Primary schools which used setting to organise their lessons in English and Mathematics, were also generally associated with negative value-added[4] subject scores locally and nationally. In brief, students who were

---

[4] In educational research value-added modeling attempts to isolate the unique contribution of schools from factors outside the school’s control that are known to strongly affect student test

taught in sets improved less in English and Mathematics. In contrast, case schools that did not adopt setting were generally associated with positive value-added scores. Their students improved more in English and Mathematics between KS1 and KS2.

Finally, a recent nationally-representative cross-sectional study of 7 year-old children in England, which used data from the MCS (Parsons and Hallam, 2014), found that children in top streams had higher KS1 scores (reading, maths and overall) compared to children in non-streamed classrooms. Children in middle and bottom streams had lower scores in their KS1 tests compared to children in non-streamed classrooms, even after controlling for child and family characteristics, school experiences at age 7, parenting behaviour and home learning environment at age 3 and 5, parent engagement with school at age 7 and school characteristics at age 7.

Most of the studies reviewed rest on the strong assumption that all the relevant confounders are adjusted for. Therefore, it is necessary to use statistical methods that can account for confounding more effectively.

2.11: Attainment grouping in UK secondary school and achievement tests

UK research in 1960's and 1970's, had highlighted an increase in academic attainment gaps between students who were taught in streams and those who were not (Gillard, 2009). Recent evidence over the last two decades, following the introduction of the National Curriculum in 1988, has produced mixed findings.

A longitudinal study of a subsample of 180 pupils in a mathematics department in a co-educational comprehensive school in Greater London, who were followed from the age of 14 to 16 years in the early 1990's, found that 'tracked' group teaching compared to mixed-attainment group teaching was beneficial for the mathematics

---

performance. In their analysis Kutnick and Colleagues (2006) adjusted for KS1 scores to explore whether academic progress between KS2 and KS1 was associated with attainment grouping.

progress of low-attaining students and had no effect for high-attaining students (Venkatakrishnan and Wiliam, 2003a).

Another longitudinal study following a cohort of 955 students between 1996 and 2000 in six secondary schools in Greater London, found that attainment grouping in mathematics had little effect on overall mathematics GSCE attainment, whereas it produced more gains in mathematics attainment for higher attaining students in the expense of lower attaining students (Wiliam and Bartholomew, 2004).

A panel study of nearly of 6000 secondary school students in 45 mixed secondary comprehensive schools (located in London, some southern counties of England to East Anglia and South Yorkshire) whose students were followed from Year 7 (1996) to Year 9 (1998-when they took the KS3[5] exams) reported that setting for English and Science had no impact on the overall KS3 attainment in English and Science even after statistical adjustment for KS2[6] exam results. Setting for Mathematics was associated with slightly higher KS3 exams scores after controlling for KS2 Mathematics score. The authors of the study also reported that pupils who attained higher levels in KS2 Mathematics tests benefited more because of setting compared to lower attaining pupils (Ireson J. et al., 2001). A follow up of the same cohort of pupils from Year 9 (1998) to Year 11 (2000) revealed that setting had no impact on overall GCSE attainment in English, Mathematics and Science when prior KS2 or KS3 scores were statistically accounted for. The findings of this study indicated that setting had no differential impact on students with higher or lower prior overall KS2 or KS3 attainment in English or Mathematics. Thus, no association was found between setting and attainment gaps in English and Mathematics.

---

[5] Secondary education is split between KS3 (age 11-14) and KS4 at age 14, to align with the two-year examination courses at GSCE level.

[6] In England and Wales, KS1 fits the first stage of primary education (age 5-7) and KS2 fits the later stage of primary education (age 8-11) which takes pupils up to standardised assessment test at age 11.

In Science, however, high-attaining KS3 students in sets achieved worse GCSE grades than high attaining students who were not set (Ireson et al., 2005).

2.12: Children's characteristics and verbal cognitive development in the UK

2.12.1: Socio-economic differences in verbal cognitive development

There are pronounced socioeconomic differences in verbal skills in the UK. These differences are already present even before children start their formal education (Dearden et al., 2011, Jerrim and Vignoles, 2013, Sindall et al., 2019). Kelly et al. (2011) used data from the MCS and found pronounced differences in children's verbal skills even from the age of 3 years. These differences widen even further between 3 and 5 years. The gap in verbal skills between the poorest children in the UK and children from privileged backgrounds, already large at age 5, grows particularly fast during primary school (Bradbury et al., 2019). Children from the poorest fifth of the family income distribution, who perform well in KS1 national exams at age 7, are more likely to fall behind their peers from the richest fifth in KS2 exams by age 11, whereas underprivileged children with a bad performance in KS1 are less likely to improve their ranking in KS2 exams than privileged children (Goodman and Gregg, 2010). The reasons for these widening socio-economic gaps in verbal skills remain unclear. What happens within family seems to matter considerably (Kelly et al., 2011, Bradbury et al., 2015), however the role of school environment in widening socio-economic gaps in the UK, particularly how schools organize the learning of their students is under researched.

2.12.2: Early verbal skills and later verbal cognitive development

Initial differences in cognitive skills, when children enter primary school in the UK , are large and remain the biggest predictor of cognitive inequalities for children on the verge of entering secondary school (Bradbury et al., 2015). Early childhood development is the first and important bottleneck in life, but it is not the only one at which inequality develops. Inequality of opportunity is apparent from the time a child is born and then accumulates throughout childhood and adolescence. These periods in one’s life coincide with formal education.

In a recent report using data from the MCS, Moss and Washbrook (2016) found that poorer performance in literacy at age 11 is strongly predicted by early language and (to a lesser extent) attention skills. In another study Zilanawala and colleagues (2017) used data from the MCS to which they applied the statistical technique of Latent Class Growth Modelling. Primary school children were divided into three verbal attainment groups (low, average, and high attaining children). The study reported that the differences between the three verbal attainment groups increased substantially between 7 and 11 years. MCS children in the high attainment group improved their verbal skills considerably, while MCS participants in the low attainment group hardly improved their verbal skills between ages 7 and 11 years.

To date, there have been no attempts to examine whether attainment grouping in primary school relates to an increase of verbal inequalities in a UK contemporary context.

2.13: Summary and limitations of past research

Contemporary evidence on the relationship between attainment grouping and achievement test scores in the UK is limited. There is also a notable paucity of scientific literature describing the association between attainment grouping and aptitude tests (verbal and non-verbal). Overall, most reviewed studies provided evidence of little or no association between attainment grouping and achievement test scores among primary and secondary aged pupils in the UK. There was also some evidence of academic benefits for high attaining students who were also grouped by attainment compared to students who were not. However, caution should be exercised when interpreting the findings of these studies.

Almost all the contemporary studies reviewed, which were published in the last two decades, did not use nationally representative data, thus making generalisations for the English context difficult.

Interestingly, no UK studies exist exploring the long-term effect of attainment grouping in primary school on later cognitive outcomes. Such studies could have generated evidence on the long-term effect of attainment grouping in primary school.

It also remains uncertain whether attainment grouping is related to an increase in socioeconomic and educational inequalities during primary school.

2.14: Literature search; Attainment grouping and psychological outcomes

Literature search, restricted in the last two decades, was conducted in Psych Info, Medline, ERIC, and Education Abstracts databases as well as manually via Google scholar. Only one non peer-reviewed study was found via Google scholar on setting for mathematics in primary school and psychological symptoms using the MCS. No other contemporary UK studies were found on attainment grouping in UK secondary school and psychological symptoms. Therefore, the rest of this section will focus on identified studies on attainment grouping and non-academic outcomes, such as liking for an academic subject or academic self-confidence.

2.15 Attainment grouping in UK primary school, liking for school and psychological symptoms

Whilst some research has been carried out on the association between attainment grouping and non-cognitive outcomes during secondary school in England, such associations in UK primary education have received scant attention in the research literature.

A qualitative study of six Year 6 students in 6 primary schools across England, who were taught in three attainment groups , found that children’s attitudes towards school were not affected by attainment grouping, but students’ awareness of their placement and the nature of teasing within school were associated, although these associations were mediated by primary school's ethos (Hallam et al., 2004b).

A recent non peer-reviewed study of almost 5000 children in English and Welsh primary schools, which used fixed effects panel data from the MCS, reported no significant differences in parent- and teacher-reported psychological symptoms between 7 and 11 years among children who were set for maths between 7 and 11 years and those who were not set for maths at both ages (McDool, 2018). This study adopted an instrumental variable approach to investigate whether placement in bottom set for Mathematics between 7 and 11 years contributed to psychological symptoms. The rationale behind the use of instrumental variable approach in this study was to account for the issue of endogeneity which is likely to arise due to reverse causality; for instance while children's behaviour may influenced by set placement, the child's behaviour is also likely to influence the set in which they are placed. Multivariable linear regression was considered as an inappropriate technique to deal with endogeneity since it would have produced biased and inconsistent estimates of the effect of placement at the bottom sets for mathematics on children's behaviour. By using two instruments; the proportion of children with English as their additional language (EAL) in child's class[7] (age 11) and the number of sets for mathematics in child's school year[8] (age 11), the author of the study

[7] The proportion of children in a child's class from homes with English as an additional language (EAL) was used as an instrument since it was assumed that these children were more likely to be overrepresented within lower sets in Mathematics in primary school, and it was also assumed that the proportion of children from EAL homes was unlikely to be associated with children's behaviour.

[8] The number of sets within the child's school year was used as an instrument since it was assumed that there was a correlation between the number of sets and placement at the bottom set; the greater the number of sets the more unlikely it was assumed for a child to be placed at a

found no evidence of a statistically significant effect of placement to the bottom set for mathematics on children's psychological symptoms.

2.16: Attainment grouping in UK secondary school, liking for an academic subject and academic self confidence

In a mixed methods study of 943 students who were followed from Year 8 to Year 9 in six secondary schools in the Greater London area, Boaler et al. (2000) found that setting for mathematics was associated with curriculum polarisation and disaffection of pupils towards mathematics. When students were taught mathematics in mixed attainment classes, their teachers gave them work that was at an appropriate level and pace. In contrast, when students were allocated to mathematics sets, the opportunity to learn was limited for students who were placed in lower sets. Students in top sets were required to learn at a faster pace which was incompatible with their understanding. As a result, students who were taught mathematics in sets were neither satisfied with their set placement, nor did they enjoy mathematics.

In another longitudinal study of 3000 students followed from Year 7 to Year 9 in 45 mixed comprehensive schools in England, Ireson et al. (2001) found that students' general academic self-concept, which is the belief of being good at school, was higher among students in the group of schools with moderate levels of setting (setting was not adopted for all core subjects such as English, Mathematics and Science) compared to groups of secondary schools with no setting or a high level of setting in English, mathematics and science (setting was adopted for all these core subjects). Setting for mathematics was not associated with the belief of being good at mathematics. Setting for English was associated with a diminished belief of being good at English for high-attaining students, while it was associated with an enhanced belief of being good at

bottom set because of the existence of more alternative sets. It was also assumed that the number of sets for Mathematics was uncorrelated to children's behaviour.

English of those students who had lower scores in KS2 exams. Caution should be exercised when interpreting these findings since the authors of the study could not control for prior belief of being good at Mathematics or English.

A study of 6000 secondary pupils from 45 secondary comprehensive schools in England found that the level of setting between Year 7 and Year 9 did not appear to be associated with pupils' liking for school in Year 9. This was true even after statistical adjustment for general academic self-concept in Year 9, prior attainment (Year 6), gender and years of teaching experience. Liking for school in Year 9 was not available, so changes in liking for school could not be explored along with changes in the level of setting. In addition, a lower score in liking for school was found among pupils who were taught in bottom sets between Year 7 to Year 9 (Ireson and Hallam, 2005).

A follow-up study of over 1600 students from 23 co-education secondary comprehensive schools found that students in schools with high level of setting in English, Maths and Science between Year 7 and Year 9, there was a greater decline in students confidence of being good at English, Maths and Science in Year 11 (Age 16 years).In addition, the number of years of exposure to setting for Maths, English or Science between Year 7 to Year 11 did not have an effect on subject-specific domains of the academic self-concept (the confidence of being good at) during Year 11. However, significant associations were found between set placement for each subject and academic self-concept for all three subjects. Students who were in top sets were more likely to have higher academic self-concept than pupils who were taught in bottom sets (Ireson and Hallam, 2009).

Finally, Francis and Colleagues (2017b) examined the association between set placement in English and Maths, subject and general self- confidence. In this mixed methods study, information was used for 11500 students who were followed between year 7 (aged 11/12) to year 8 (12/13) in 139 secondary schools in England. By comparing pupils in sets with pupils in mixed-attainment classes for English and Maths, the study found a significant positive relationship between perceived set placement, subject confidence and general confidence; those students in lower sets reported lower levels of both subject and general confidence. The authors of the study argued that the practice of setting in English secondary schools labelled students according to ability and was responsible for a self-fulfilling prophecy whereby students behaved in par with the label of their set placement.

2.17: Children's characteristics and psychological symptoms in the UK

Although the lack of published studies on the relationship between attainment grouping in UK schools and socioemotional difficulties, there is ample empirical evidence on the association between socio-economic circumstances, school readiness and externalising and internalising symptoms.

2.17.1: Socio-economic circumstances and children's psychological symptoms

A considerable amount of literature, using mainly US data, has been published on the association between family income and psychological symptoms during childhood and adolescence (Cooper and Stewart, 2013a, Cooper and Stewart, 2017). The greater part of the published literature on the issue has reported higher levels of psychological symptoms for children and adolescents from families with low income. Besides, there is evidence that children and adolescents from lower income families tend to benefit more when the income of their families increases compared to children and adolescents from higher income families.

In the UK, stark income gradients in psychological symptoms are observed from a very early age in childhood (Kiernan and Huerta, 2008, Kiernan and Mensah, 2009, Kelly et al., 2011), which are sustained through adolescence (Borrell-Porta et al., 2017, Noonan et al., 2018). The mechanisms through which family income exerts an influence on children's and adolescents' psychological symptoms have been explored through psychological and economic theoretical frameworks, particularly the parental stress and parental investment theories. The parental investment theory reflects the ability of parents to invest in resources, whether they are material or emotional, in order to support their children's wellbeing, as well as their learning and cognition (Duncan et al., 2014). The family stress theory suggests that adverse social and economic circumstances, such as poverty, may create a stressful atmosphere within the home environment, which in turn may affect the psychological wellbeing of parents and subsequently their parenting style and practices (Cooper, 2017).

The Education Endowment Foundation (2018a) highlights a paucity of studies exploring the interaction between attainment grouping, familial socio-economic circumstances and children's psychological symptoms .

2.17.2: Verbal skills and children's psychological symptoms

There is some evidence that verbal cognitive ability is associated with psychological symptoms over the life course (Flouri et al., 2014, Flouri et al., 2012, Schoon et al., 2010, Zilanawala et al., 2017, Petersen et al., 2013, Hurry et al., 2018). It has also been argued that higher verbal skills may protect against childhood adversity, thus promoting resilience in socioemotional development. For instance, Flouri and colleagues (2012) (2014) found that higher verbal skills may be protective against the detrimental influence of neighbourhood deprivation or family poverty on internalising symptoms in childhood using the data from the MCS. Another recent study which used the MCS data, showed that children who had consistently below average verbal skills across childhood, were at increased risk of higher psychological symptoms at age 11 than children with consistent above average verbal skills (Zilanawala et al., 2017). Results from the 1970 British Cohort Study (BCS70), following nearly 7000 men and women who were born in the UK in 1970, showed that language skills at age 5 were significantly associated with better mental health outcomes in adulthood, even after statistical control for family background and experiences of social adaptation (Schoon et al., 2010).

What is not known yet and it should then be explored is whether attainment grouping exacerbates socioeconomic differences in children's psychological symptoms, or whether it increases the gap in psychological symptoms between high and low attaining children.

2.18: Summary of findings and limitations of past UK research

Given the passions that are aroused whenever there is a debate either for the selection of pupils between schools (comprehensive vs. grammar school or between fee-paying vs. state-maintained schools) or selection within schools (attainment grouping vs. mixed-attainment), there is a surprising lack of empirical research in the UK contemporary context comparing the social and emotional functioning of pupils who are taught in attainment groups compared to those who are not taught in streamed or set classrooms.

Overall, the past literature is suggestive of a negative association between attainment grouping and academic self-concept or enjoyment of an academic subject, while there is some evidence of heterogeneous effects by group level placement. Students who are allocated to bottom groups are more likely to experience a decrease in their academic self-concept over time compared to students who are allocated to the top groups. However, evidence remains narrow in focus dealing mostly with the association between attainment grouping and non-cognitive outcomes of secondary school pupils. Besides, there is a paucity of UK studies exploring the long-term effect of attainment grouping on children's social and emotional outcomes from primary to secondary school. Such studies could have contributed to generating evidence on the long-term impact of attainment grouping in primary school on children's externalising and internalising symptoms.

The generalisability of most published findings to the UK context is also problematic. Most reviewed studies used non-representative data from England, with only one study using nationally representative data for England and Wales.

The investigation of the effect of streaming or setting in isolation, as most reviewed studies did, is difficult when different streaming and setting policies operate in the same school simultaneously. It would be then preferable to explore the combined effect of streaming and setting, as these grouping practices are usually combined in the evidence reviews on the effect of attainment grouping (Education Endowment Foundation, 2018a).

The existing literature on the association between attainment grouping and non-cognitive outcomes has not considered the possibility of attainment grouping exacerbating socio-economic and initial attainment inequalities.

Finally, the existing literature has yet to address the methodological issues arising from non-statistical control for hard to measure time-invariant factors that could influence verbal growth and psychological symptoms over time. It also remains problematic that the applied statistical approaches were not able to rule out time-varying confounding.

Chapter 3: Study aims, objectives and contributions

The general aim of this PhD is twofold: to investigate whether attainment grouping in primary school is associated with verbal skills and externalising and internalising symptoms. It will also be explored whether attainment grouping in primary school exacerbates socioeconomic inequalities in verbal skills as well as psychological symptoms, while it will be investigated whether attainment grouping during primary school is associated with an increase in socio-economic inequalities in psychological symptoms or an increase in psychological symptoms among MCS participants with different levels of verbal skills at age 5.

The focus of this PhD is placed upon the effects of between-class attainment grouping, either streaming or/and setting on children's developmental outcomes than the effect of being in a set/stream (top, middle, bottom). Thereafter, the term attainment grouping refers to between-class attainment grouping.

The decision to adopt an attainment grouping method is likely to depend on the decisions of schools' senior leadership teams about the most effective way of educating students in their schools. It is therefore plausible that the decision of school decision makers to adopt attainment grouping is largely unaffected by children's background. In contrast, allocation to different attainment groups (top, middle, bottom) by teachers may be dependent on children's social background, educational disadvantage (e.g. attending an under-performing school), educational advantage (e.g. having highly aspirational parents or employing tutors), variable rates of development or "birthday" effects, and therefore may be biased (McManus, 2010). In this thesis streaming and setting, two otherwise variable types of attainment grouping in terms of stringency (see subsection 2.2.), are combined together as primary schools tend to exercise them in tandem, as shown later in chapter 4 (see subsection 4.3.1.4).

Chapters 5 and 6 explore the association between the transition to attainment grouping (yes/no) between MCS 4 (age 7) and MCS 5 (age 11) and changes in children's verbal skills and psychological symptoms during the same period which coincides with KS2 in English and Welsh primary school. The biggest advantage of using a transition variable between MCS waves is facilitating a comparison in the change of children's developmental outcomes among two groups of children who were not grouped by attainment (streaming/setting) at age 7, but one of them changed to attainment grouping between 7 and 11 years. Chapter 7 explores the pooled association between attainment grouping at age 7 and trajectories of psychological symptoms from age 7 (MCS 4) to age 14 (MCS 6) using the English, Welsh, Scottish and Northern Irish MCS samples.

A description of specific aims, objectives and related hypotheses is given below followed by an account of possible contributions of this PhD on the debate around attainment grouping.

3.1: Aims, Objectives and Hypotheses

Aim 1: **To examine whether a transition from mixed attainment grouping to attainment grouping during KS2 is related to a change in children's verbal skills or an increase in inequality over the same period.**

Objective 1.1: To examine whether a transition from mixed attainment grouping to an attainment grouping (streaming/setting for literacy) between ages 7 and 11 years is related to progress in children's verbal skills.

Hypothesis 1.1: MCS participants who transitioned from mixed attainment grouping to attainment grouping (streaming/setting for literacy) between ages 7 and 11 years will exhibit slower improvement in verbal skills compared to children who were taught in mixed attainment groups at both ages 7 and 11.

Objective 1.2: To test whether a transition from mixed to attainment grouping between ages 7 and 11 years is related to a change in socioeconomic inequalities in children's verbal skills between ages 7 and 11 years.

Hypothesis 1.2: A transition from mixed to attainment grouping (streaming/setting for literacy) between ages 7 and 11 years will be associated with an increase in socio-economic inequalities in verbal skills from age 7 to age 11. The verbal skills gap between the most advantaged and the least advantaged MCS participants will be wider by the end of primary school among MCS participants who transitioned from mixed to attainment grouping. The gap in verbal skills between the most and the least advantaged MCS participants who are taught in mixed attainment groups at both ages 7 and 11 years will increase by less than those who changed to attainment grouping.

Objective 1.3: To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a change in the gap in verbal skills between age 7 and 11 contingent on the gap in verbal skills at age 5.

Hypothesis 1.3: Contingent on the gap in verbal skills at age 5 (high versus low scores), the gap in verbal skills will widen for children who transition from mixed to attainment groups between age 7 and 11 compared to those who remain in mixed groups.

Aim 2: **To examine whether a transition from mixed attainment grouping to attainment grouping (streaming/setting) during KS2 is related to a less favourable change in children's externalising or internalising symptoms over the same period, and an increase in socioeconomic and educational-related inequalities in children's psychological symptoms.**

Objective 2.1: To examine whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a less favourable change in children's externalising or internalising psychological symptoms from age 7 to age 11.

Hypothesis 2.1: Being exposed to a transition from mixed grouping to attainment grouping between ages 7 and 11 years will be associated with a smaller decrease in parent- or teacher-reported externalising psychological symptoms and with a bigger increase in parent- or teacher-reported internalising psychological symptoms from age 7 to age 11.

Objective 2.2: To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is related to an increase in socioeconomic inequalities in children’s externalising and internalising psychological symptoms between ages 7 and 11 years.

Hypothesis 2.2: A transition from mixed to attainment grouping between ages 7 and 11 years will be associated with an increase in socio-economic inequalities in externalising and internalising psychological symptoms. The gap in externalising and internalising psychological symptoms between the most advantaged and the least advantaged MCS participants will be wider by the end of primary school among MCS participants who transitioned from mixed to attainment grouping. The gap in externalising and internalising psychological symptoms from 7 to 11 years among the most and the least advantaged MCS participants in mixed attainment groups will not increase as for those in attainment groups.

Objective 2.3: To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with an increased gap in externalising and internalising psychological symptoms between MCS participants with higher verbal skills and participants with lower verbal skills at age 5.

Hypothesis 2.3: The gap in externalising and internalising psychological symptoms for MCS participants with different levels of verbal skills at age 5 (high versus low scores) will widen among those who transition from mixed to attainment groups between age 7 and 11 compared to those who remain in mixed groups.

Aim 3: **To examine the long-term role of attainment grouping (streaming or setting) at age 7 on children’s psychological symptoms up to age 14.**

Objective 3.1: To examine whether attainment grouping at age 7 is associated with a less favourable change in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.

from 7 to 14 years, while mixed attainment grouping at age 7 is associated with a more favourable change in externalising and internalising psychological symptoms.

Hypothesis 3.1: Being exposed to attainment grouping at age 7 will be associated with a slower decrease in parent-reported externalising and a faster increase in parent-reported internalising psychological symptoms between ages 7 and 14 years, compared to mixed attainment grouping at age 7 which will be associated with a faster decrease in externalising psychological symptoms and a slower increase in internalising psychological symptoms between ages 7 and 14 years.

Objective 3.2: To test whether attainment grouping at age 7 is related to an increase in socioeconomic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.

Hypothesis 3.2: Attainment grouping at age 7 will be associated with an increase in socio-economic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years. Mixed grouping at age 7 will be associated with a smaller increase in socio-economic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.

Objective 3.3: To test whether attainment grouping at age 7 is associated with an increase in differences in externalising and internalising psychological symptoms between 7 and 14 years among MCS participants with higher and lower levels of verbal skills at age 5.

Hypothesis 3.3: The gap in externalising and internalising psychological symptoms will increase between age 7 and 14, contingent on baseline differences in verbal skills (age 5) for those MCS participants in attainment groups at age 7. The gap in externalising and internalising psychological symptoms among MCS participants in mixed-attainment groups with higher and lower levels of verbal skills at age 5 will not increase as for those in attainment groups.

3.2: Contributions of this research

This research focuses on the between-class attainment grouping practices (streaming or/and setting) in primary school; despite the common implementation of both setting and streaming in UK primary schools, few studies in the past have focused the combined effect of attainment grouping on children of this age. Attainment grouping has become common practice in UK primary schools over the last two decades, however most published studies have focused on attainment grouping and student outcomes in secondary school. This is because of the belief that attainment grouping is not that common in primary schools. Therefore, little, if anything, is known about the attainment grouping in children's cognitive and psychological outcomes in primary school.

This research contributes to the limited literature on the effects of attainment grouping in the UK (Education Endowment Foundation, 2018a) by addressing the methodological issues of unobserved time invariant and time-varying confounding surrounding the measurement of the effect of attainment grouping during primary school. Specifically, it explores how changes in attainment grouping associate with changes in children's verbal skills and psychological symptoms during primary school. As it was mentioned earlier in the chapter, a transition to attainment grouping during primary school is believed to be plausibly unconfounded by individual and family-level characteristics. Some primary schools in the UK choose to resort to attainment grouping during Key Stage 2 for reasons not related to individual and family characteristics to raise academic standards. It can then be argued that the effect of attainment grouping on developmental outcomes is not biased by individual and family level characteristics.

In addition, this research contributes to the existing literature by investigating whether attainment grouping associates with an increase in socio-economic and educational inequalities in children's developmental outcomes during primary school. To the best of my knowledge, no UK studies exist with a specific focus on attainment grouping in exacerbating well-documented socio-economic and attainment inequalities in verbal skills and psychological symptoms over time.

Besides, using verbal skills (an aptitude test outcome) is also novel to this area of literature which, to date, has focused on standardized national achievement test scores such as KS scores (a performance test outcome).

Very few published studies in the existing literature have attempted to examine the long-term relationship between attainment grouping at age 7 and the development of psychological symptoms from childhood to early adolescence (up to age 14 years) in a UK contemporary and nationally representative context. It is hoped that this research will contribute to a deeper understanding of whether the trajectories of MCS participants who were grouped by attainment at age 7 changed considerably during the transition from childhood to adolescence. Finally, studying the early adolescent period has the potential to convey more about the long-term effects of attainment grouping in primary school.

Chapter 4: Data, aims, objectives, hypotheses. and methods

4.1: Chapter Overview

The purpose of this chapter is to describe and discuss the data and the methods used in this PhD. This chapter provides a brief overview of the MCS on which quantitative analysis was conducted. This chapter also includes information on the MCS sample design, the structure of data collection, emergent data-handling issues as well as information on accessing the MCS data and information on ethical approval of data collection. A detailed description of the measures and the methods used in the results chapters is also given.

4.2: The Millennium Cohort Study

The MCS was established to capture the lives of children living and growing up in each of the four countries of the UK in the new century. The MCS is a rich dataset that provides data about children's family circumstances and the broader socio-economic context in which the children grow up. The data include measures about cognitive and socioemotional development, measures about the socio-economic circumstances in which children and adolescents live and grow, such as family income, and a number of school-related measures such as the type of school they attend, school structural information and even information about teaching practices which the MCS cohort member experienced during his/her primary school years. Therefore, the wealth of information available in the MCS makes it a suitable source for the study of the potential effects of attainment grouping practices on children's development through primary school and the transition to secondary school, as well as the possible role of such practices in cognitive and psychological inequalities apparent in the UK (Viner et al., 2018, Chzhen et al., 2018, OECD, 2018a).

At the time of writing, data were available for the first six sweeps of the study (MCS1 – MCS6), collected when the children were 9 months, 3 years, 5 years, 7 years, 11 and 14 years, which correspond to years 2001, 2004, 2006, 2008, 2012 and 2015. All analyses used information from MCS3 to MCS6, which correspond to ages spanning from 5 to 14 years.

4.2.1: MCS sample design

The MCS sample was drawn from all live births in the UK over 12 months from September 1st in England and Wales and December 1st in Scotland and Northern Ireland in the year 2000 who survived to 9 months. A random sample of electoral wards was disproportionately stratified to safeguard sufficient representation of all four UK countries, deprived areas and areas with high numbers of Black and Asian families (Hansen et al., 2014).

There are nine strata from which electoral wards were sampled at MCS1. A summary of the MCS design strata, number of wards sampled and final sample size in given in Table 4.1. A more detailed account on the MCS sample design can be found in the "*Technical report on sampling"* where the numbers in the table are reproduced from (Plewis, 2007). Wards were classified as "disadvantaged" if they were among the poorest 25% of wards in 1998, based on the Child Poverty Index (CPI), and "ethnic" (applicable in England only) if at least 30% of the population were Black or Asian.

Table 4. 1:Summary of MCS strata design

Country	Stratum	Number of Wards in stratum	Final sample size
England	Advantaged	110	4,678
	Disadvantaged	71	4,592
	Ethnic	19	2,416
Wales	Advantaged	23	844
	Disadvantaged	50	1,954
Scotland	Advantaged	32	1,163
	Disadvantaged	30	1,207
Northern Ireland	Advantaged	23	735
	Disadvantaged	40	1,120
Total		398	18,818

4.2.2: MCS sample sizes and attrition

The achieved sample size at MCS 1 was 18,552 households, which corresponds to an overall response rate of 72%. At MCS 2, families who had been eligible to participate in the MCS 1 but could not be contacted during MCS 1 were invited to enter the study. As a result, 692 new families joined the study at MCS 2 (Hansen et al., 2014). The achieved sample sizes for all MCS sweeps are shown in Table 4.2.

Table 4. 2: Achieved sample sizes for MCS sweeps 1 to 6 (all UK countries)

Sweep	N (children)	N (families)
MCS 1	18,818	18,552
MCS 2	15,808	15,590
MCS 3	15,459	15,246
MCS 4	14,043	13,857
MCS 5	13,469	13,287
MCS 6	11,726	11,576

4.2.3: Structure of data collection

For each sweep, an interviewer would visit the child's home to conduct the interview. When the cohort member was younger, more questions were asked to the main carer (mainly biological mothers of the MCS participants). However, as the MCS children got older, they were directly asked more questions. At each sweep, there were three different versions of parent interviews that could be completed; these were: parent, partner, and partner proxy. The person who answered the questions depended on the composition of the household. In most cases any parents of cohort members and partners of parents were interviewed. If no parents were present, the child's main carer was interviewed. If the person selected for the partner interview was not present at the time of the interview, the main carer would complete the 'partner proxy' interview on behalf of their partner. A mixture of self-completion surveys and face-to-face interviews were used to collect the information from cohort members (CMs) and their parents. The interviewer followed a schedule using a computer-assisted personal interviewing programme (CAPI) which automatically routed the questions depending on the answers given. Respondents in Wales were provided with all main communication materials in English and Welsh and were also able to choose which language they participated in (Fitzsimons et al., 2017a). In order to support participation of parents with limited proficiency in English, other language materials were provided in seven languages most required during MCS sweeps: Arabic, Bengali, Gujarati, Hindi, Punjabi (Gurmukhi script), Punjabi (Urdu script) and Urdu. When the main and partner respondents were unable to speak English or were uncomfortable with completing the interview in English, interviewers were instructed to find an interpreter or if an interpreter could not be found a bilingual interviewer was reallocated to the address (Fitzsimons et al., 2017a).

MCS questionnaires were also delivered to teachers at MCS 4 and MCS 5. Teachers were asked questions about children's behaviour in the classroom, their attainment in different subjects, the composition of their classrooms, their teaching practices, their teaching qualifications and the number of years in the teaching profession

(Johnson et al., 2011). MCS cohort members were also asked to fill in questionnaires at MCS 4, MCS 5 and MCS 6, with questions covering a wide range of topics from children's interests, school-related questions and about their mental health. Child cognitive assessments were available in the MCS; from MCS 2 to MCS 6. Figure 4.1 gives an overview of the survey content at each MCS sweep.

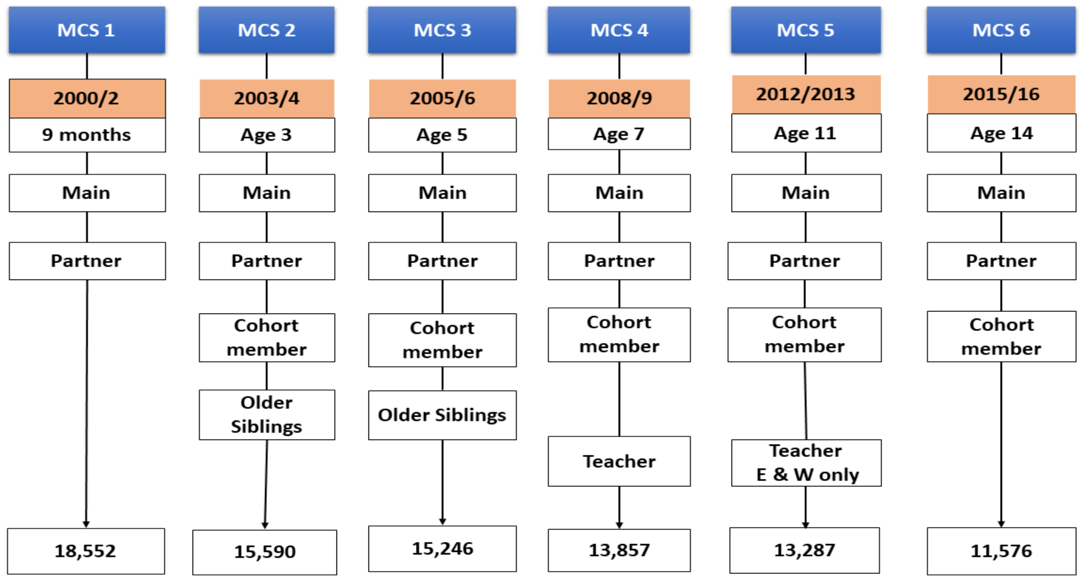

Figure 4. 1MCS survey content modified from (Hansen et al., 2014) and (Fitzsimons et al., 2017a)

4.2.4: MCS teacher surveys

Teacher surveys were administered at MCS 4 and MCS 5. The MCS4 teacher survey included responses of MCS children's teachers from the 4 countries of the UK, while the MCS 5 teacher survey was administered only in England and Wales due to funding constraints in Scotland and Northern Ireland. Teacher responses at both MCS 4 and MCS 5 were used in the statistical analyses of chapter 5 and chapter 6 in order to explore the association between a change to attainment grouping between 7 and 11 years and a change in verbal skills, externalising and internalising psychological symptoms between 7 and 11 years. Verbal skills using the British Ability Scales (BAS) were measured up to age 11, so it was decided to explore the long term association between attainment grouping at age 7 years (early attainment grouping) and trajectories of externalising and internalising symptoms from 7 to 14 years of age in the UK, for which parent responses for the Strengths and Difficulties Questionnaire (SDQ) were available. What follows is a brief description of the achieved sample in teacher surveys at MCS 4 and MCS 5, an account of the patterns of teacher non-response at both MCS waves, and a comparison between the achieved MCS sample in England and Wales at both MCS 4 and MCS 5 and teacher responses at the same period.

4.2.4.1: Teacher survey at MCS4 (England, Wales, Scotland, and Northern Ireland)

In the year 2008, questionnaires were sent to MCS child's teacher. The main MCS4 sample consisted of 13,682 families and 13,857 cohort members. From a total of 13,857 cohort members, 5,382 teachers from UK primary schools answered and returned questionnaires for 8,775 seven-year old (63.2%), and 5082 (36.8%) had completely missing records.

4.2.4.2: Teacher Survey at MCS5 (England and Wales only)

The MCS 5 teacher survey took place in 2012 in England and Wales only due to funding constraints in Scotland and Northern Ireland (Gallop et al., 2013) From a total of 9,610 eligible MCS participants in England and Wales, 5,079 primary school teachers in England and Wales answered and returned questionnaires for 7,430 MCS participants (77.3%), and 1,602 (22.7%) had completely missing records.

4.2.4.3: Patterns of teacher response at MCS4 & MCS5 (England and Wales only)

Out of 9575 eligible MCS participants who were eligible for participation in the MCS teacher surveys in England and Wales at both MCS 4 and MCS 5, teachers returned questionnaires for 46.7% (N=4469) of MCS children at both occasions, while information was provided only at MCS 4 for 23.8% of children and for 29.5% only at MCS 5 (see Table 4.3).

Table 4. 3: Patterns of teacher response at MCS sweeps 4 and 5 (N=9,575)

Pattern	Percent (%)	N
Teachers returned questionnaires at both MCS sweeps	46.7%	4,469
Teachers returned questionnaires at MCS 4, but not at MCS 5	23.8%	2,282
Teachers returned questionnaires at MCS 5, but not at MCS 4	29.5%	2,824

4.2.4.4: Comparison between whole sample and sample with teacher response at MCS sweeps 4 and 5

Analyses were undertaken in order to investigate how representative were MCS participants whose teachers responded to both MCS4 and MCS5 compared to the whole sample of MCS participants who were eligible for participation in the MCS teacher surveys at both MCS sweeps in England and Wales. First, sample composition statistics as in Table A1.1 in Appendix 1 were computed with two different samples; the entire eligible MCS sample in England and Wales at MCS 4 and MCS 5 including twins and triplets (N=9,575) and the sample with responses to teacher questionnaire at MCS 4 and MCS 5 (N=4,469).

Response models used information from all variables from Table A1.4 in Appendix 1. Given the varying probabilities of selection inherent in the MCS sample design, analysis of MCS data takes account of the design sampling weights and weights adjusting for differential non-response across the various sweeps. The dependent variable was modelled as binary and it was given the value of 1 if the teachers of the MCS child responded to both MCS 4 and MCS 5 and 0 otherwise.

Table A1.2 shows that the results were not very different if the distributions were estimated only for the teacher survey sample compared to the whole eligible MCS sample in England and Wales. When those children whose teachers responded to the teacher survey were considered, it was found that there were more MCS children from advantaged areas in England, there were more two-parent, dual-earner families, English-only speakers, ethnically White and higher income families in comparison with the entire MCS4 and MCS5 sample in England and Wales. The differences in the two considered samples were small since differences were smaller than 6%. This highlights that teacher non-response to both surveys (53.3%) did not associate with great differences in the composition of the two considered samples. This finding was further confirmed by multivariable logistic regressions shown in Table A1.2.

Table A1.2 shows that the estimates were mostly not statistically significant at the conventional $p<0.05$ level except for the MCS participants from ethnically dense areas in England, those from low-income families and Bangladeshi ethnic background. These variables were related to a lower likelihood of teacher response at both MCS sweeps. In contrast, those MCS participants with at least one parent who was working had a higher likelihood of teacher response at both MCS sweeps compared to children whose parents were both working. These findings were indicative of a weak statistical relationship between MCS participants' characteristics and the probability of teachers responding. In addition, the low pseudo R-squared indicated that the predictive power of the overall model was weak.

The descriptive statistics and the multivariable logistic regression models highlighted that teacher non-response did not dramatically shift the distribution of the analytic sample and that teacher response was only weakly explained by child's characteristics. It is worth stating that provided the available MCS data, it was not possible to explore whether school-level characteristics could explain teacher non-response at both MCS sweeps. Non-response to both MCS teacher surveys was assumed to be MCAR since previous analyses by Mostafa and Rosenberg (2013) showed that non-response to MCS4 teacher survey in England was Missing Completely At Random (for a detailed description of missing data mechanisms see section 4.6.1).

4.2.5: Data structures and data-handling issues

The data structure format available for MCS 1 to MCS 4 was wide. However, the format of the MCS data changed from wide to long at MCS 5 so as to combine data longitudinally more straightforwardly (Fitzsimons et al., 2017a). In order to facilitate the combination of information from MCS 5 & MCS 6 and MCS 3 & MCS 4, several variables of interest were selected from MCS 5 and MCS 6 and they were reshaped from long to wide format by using the person identifiers (EPNUM00 and FPNUM00). When all the MCS data were in wide format the unique family identifier (MCSID) was used to combine data from MCS 3 to MCS 6.

4.2.6: Access to MCS data

The data sources used were available for internet download from the UK Data Service at the University of Essex. After signing the relevant confidentiality data agreements, datasets were downloaded together with the relevant documentation. The persistent identifiers for MCS datasets are:

- MCS3 survey (2006) http://doi.org/10.5255/UKDA-SN-5795-4
- MCS4 survey (2008) http://doi.org/10.5255/UKDA-SN-6411-7
- MCS4 teacher survey (2008) http://doi.org/10.5255/UKDA-SN-6848-1
- MCS5 survey (2012) http://doi.org/10.5255/UKDA-SN-7464-4 (it includes MCS5 teacher survey)
- MCS6 survey (2015) http://doi.org/10.5255/UKDA-SN-8156-4

4.2.7: Ethical considerations

Ethical approval for the MCS was obtained from the relevant ethics committees at each wave of data collection (Multi-Centre Research Ethics Committee (MREC) and National Health Service Ethical Authority) and parents gave informed consent before interviews (Hansen et al., 2014, Fitzsimons et al., 2017a). Further approvals were sought and given for carrying out the survey in schools in each UK country. For England, the survey was approved by the Star Chamber in the Department for Children, Schools and Families; for Wales, by the Schools Workforce Advisory Panel; for Scotland, by the Directors of Education in the Local Educational Authorities; and for Northern Ireland, no formal approval was needed (Gallop et al., 2013, Huang and Gatenby, 2010).

4.3: Measures

4.3.1: Attainment Grouping

The information on streaming or setting was provided by the MCS participant's teacher who was asked the following questions at MCS 4 (age 7) and MCS 5 (age 11) "I*n this child's year is there streaming?", "In this child's year are there sets for literacy?"* and *"In this child's year are there sets for maths?"* Within the teacher questionnaire detailed definitions of streaming, setting or within-class grouping were given, thus reducing the possibility of teachers conflating between-class grouping (streaming or setting) with within-class grouping. It should be noted that children who were not taught in streams or sets might still have been taught in groups for specific subjects within the classroom.

Teachers who gave an affirmative answer to the above-mentioned questions were further asked at MCS 4 and MCS 5 *"Which stream is this child in?"*, *"Which set is this child in for literacy?"* and *"Which set is this child in for maths?* What follows is an analysis and a discussion of the prevalence of attainment grouping and the characteristics of the English and Welsh primary schools which adopted streaming and setting at MCS 4 and MCS 5 (a detailed description of the available MCS school-related characteristics used for the analyses of this chapter is given in subsection 4.3.4).

Cross tabulations were conducted to explore the distribution of MCS participants' characteristics in the top, middle and bottom groups, followed by pooled univariable multinomial logistic regressions at MCS 4 and MCS 5 to quantify the probability of children's membership either in the top, middle or bottom groups (streams or sets) by their individual and family characteristics (a detailed account of the measures used in section 4.3.1.3 can be found in Table A1.3 in Appendix 1). Because observations at MCS 4 and MCS 5 are clustered within MCS participants, analyses were adjusted for the nonindependence of observations at MCS 4 and MCS 5 using the CLUSTER command in Stata. This procedure is derived from the Huber-White variance estimator and provides robust standard errors adjusted for within-cluster correlated data (Rogers, 1994).

Finally, a detailed account of the relationships between different types of attainment grouping and group membership is presented, followed by a description of transitions to attainment grouping relative to available MCS school related characteristics.

4.3.1.1: Prevalence of streaming and setting at MCS 4 and MCS 5

During the MCS 4, 17.5% of MCS participants were taught in streams, with the percentage increasing to 21.3% at MCS 5. Setting for maths was more prevalent than setting for literacy at both MCS 4 and MCS 5. At MCS 4, 29.6% of children were taught in sets for literacy and 42.2% at MCS 5. Setting for maths was the most prevalent type of attainment grouping at MCS 4 and MCS 5. At MCS 4, 34.7% of MCS participants were taught in sets for maths, while 64.3% were taught in sets for maths at MCS 5.

4.3.1.2: Prevalence of streaming and setting at MCS 4 and MCS 5 by school related characteristics

Table 4.4 highlights that 21.2% at MCS 4 and 25.4% of MCS participants at MCS 5 who were attending larger primary schools (highest tertile of the number of classes in child’s year) were taught in streamed classrooms compared to 13.1% and 17.1% of MCS participants who were attending smaller primary schools in England and Wales (lowest tertile of the number of classes in child’s year) at MCS 4 and MCS 5 respectively. In addition, primary schools in England and Wales serving more diverse student populations in terms of the number of children with SEN and EAL were more likely to adopt streaming at MCS 4 and MCS 5. Between MCS 4 and MCS 5, the biggest increase in adopting streaming was observed among fee-paying primary schools. At MCS 4, 9% of MCS participants were attending fee paying primary schools that were adopting streaming, while the percentage of MCS participants attending fee-paying primary schools which were adopting streaming tripled to 27.5% at MCS 5. Notable increases between MCS 4 and MCS 5 were evident for MCS participants who were being taught in classrooms with a high number of classmates with English as an Additional Language (EAL) and by teachers with many years of teaching experience. Tables 4.5 and 4.6 show that larger primary schools and primary schools in which MCS participants were taught in the same classroom with a higher number of MCS participants from an EAL background were more likely to adopt streaming for literacy and maths. The biggest percentage increase between MCS 4 and MCS 5 in setting for literacy, but not as substantial as streaming, was observed for fee-paying primary schools (a 19.3% increase). For maths literacy the biggest increase was seen for larger primary schools, followed by fee-paying primary schools.

Table 4. 4: Prevalence of streaming at MCS 4 and MCS5 by school-related characteristics.

	Streaming at MCS 4 (Age 7)			Streaming at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Type of school						
State-maintained	19.1%	17.3%	20.8%	21.4%	19.3%	23.4%
Fee-paying	9.0%	4.4%	13.5%	27.5%	21.5%	33.5%
Faith school	14.8%	12.0%	17.7%	18.2%	15.3%	21.1%
No disruptive peers in child's classroom	17.0%	15.1%	18.9%	21.0%	18.9%	23.0%
Disruptive peers in child's classroom	18.4%	15.8%	21.0%	22.0%	19.6%	24.5%
Number of classes in child's year						
Lowest tertile (Mean=1)	13.1%	10.9%	15.4%	17.1%	15.8%	18.5%
Medium tertile (Mean=2)	19.3%	16.7%	21.8%	22.8%	20.9%	24.8%
Highest tertile (Mean=3)	21.2%	17.7%	24.7%	25.4%	23.5%	27.4%
Number of children in child's class						

	Streaming at MCS 4 (Age 7)			Streaming at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Lowest tertile (Mean=20	17.0%	14.5%	19.6%	21.6%	18.9%	24.2%
Medium tertile (Mean=27)	17.9%	15.4%	20.4%	21.0%	18.5%	23.6%
Highest tertile (Mean=30)	17.5%	14.6%	20.4%	18.6%	15.0%	22.2%
Child in single-year group	16.7%	14.8%	18.5%	20.4%	18.6%	22.2%
Child in mixed-year group	20.5%	17.1%	24.0%	24.3%	21.1%	27.4%
Number of SEN children in child's class						
Lowest tertile (Mean=0)	15.1%	12.9%	17.2%	18.2%	15.9%	20.4%
Medium tertile (Mean=1)	19.3%	16.1%	22.5%	21.4%	18.9%	24.0%
Highest tertile (Mean=5)	20.3%	17.3%	23.2%	27.6%	24.1%	31.2%
Number of EAL children in child's class						
Lowest tertile (Mean=0)	17.3%	15.1%	19.6%	17.3%	15.1%	19.5%
Medium tertile (Mean=1)	17.6%	14.8%	20.4%	22.5%	19.2%	25.7%

	Streaming at MCS 4 (Age 7)			Streaming at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Highest tertile (Mean=8)	17.6%	14.6%	20.5%	25.0%	21.6%	28.3%
Years of teaching experience						
Lowest tertile (Mean=3.8 years)	21.2%	18.2%	24.3%	23.4%	20.6%	26.1%
Medium tertile (Mean=12.6 years)	17.3%	15.0%	19.7%	19.7%	17.1%	22.3%
Highest tertile (Mean=22.7 years)	13.3%	10.8%	15.8%	20.4%	17.4%	23.3%

Table 4. 5: Prevalence of literacy setting at MCS 4 and MCS 5 by school-related characteristics.

	Literacy setting at MCS 4 (Age 7)			Literacy setting at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Type of school						
State-maintained (Ref.)	30.9%	28.4%	33.4%	43.3%	40.0%	46.7%
Fee-paying	23.9%	16.3%	31.6%	43.3%	35.5%	51.0%
Faith school	27.0%	23.4%	30.7%	39.0%	34.8%	43.3%
No disruptive peers in child's classroom	28.9%	26.6%	31.3%	42.4%	38.9%	45.8%
Disruptive peers in child's classroom	30.9%	27.4%	34.4%	41.9%	38.1%	45.6%
Number of classes in child's year						
Lowest tertile (Mean=1)	23.6%	20.6%	26.5%	31.6%	28.1%	35.0%
Medium tertile (Mean=2)	32.6%	29.4%	35.8%	48.7%	44.4%	53.1%
Highest tertile (Mean=3)	33.9%	29.2%	38.7%	47.8%	41.4%	54.3%
Number of children in child's class						
Lowest tertile (Mean=20)	26.8%	23.8%	29.8%	38.2%	34.7%	41.7%

	Literacy setting at MCS 4 (Age 7)			Literacy setting at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Medium tertile (Mean=27)	29.2%	26.2%	32.2%	46.3%	41.8%	50.7%
Highest tertile (Mean=30)	33.6%	29.7%	37.6%	40.7%	34.5%	46.8%
Child in single-year group	29.7%	27.3%	32.1%	42.6%	39.1%	46.1%
Child in mixed-year group	29.1%	25.3%	33.0%	40.9%	36.9%	44.9%
Number of SEN children in child's class						
Lowest tertile (Mean=0)	27.7%	25.0%	30.4%	40.2%	36.4%	44.0%
Medium tertile (Mean=1)	31.1%	26.9%	35.3%	42.3%	38.6%	46.0%
Highest tertile (Mean=5)	31.6%	28.0%	35.2%	45.9%	41.5%	50.3%
Number of EAL children in child's class						
Lowest tertile (Mean=0)	26.6%	23.8%	29.4%	39.7%	36.2%	43.3%
Medium tertile (Mean=1)	29.8%	26.4%	33.3%	38.6%	34.0%	43.2%
Highest tertile (Mean=8)	34.3%	30.5%	38.2%	49.7%	44.8%	54.7%

	Literacy setting at MCS 4 (Age 7)			Literacy setting at MCS 5 (Age 11)		
	%	95% CI	95% CI	%	95% CI	95% CI
Years of teaching experience						
Lowest tertile (Mean=3.8 years)	26.8%	23.8%	29.8%	38.2%	34.7%	41.7%
Medium tertile (Mean=12.6 years)	29.2%	26.2%	32.2%	46.3%	41.8%	50.7%
Highest tertile (Mean=22.7 years)	33.6%	29.7%	37.6%	40.7%	34.5%	46.8%

Table 4. 6: Prevalence of setting for maths at MCS 4 and MCS5 by school-related characteristics

	Numeracy setting at MCS 4 (Age 7)			Numeracy setting at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Type of school						
State-maintained	30.9%	28.4%	33.4%	43.3%	40.0%	46.7%
Fee-paying	23.9%	16.3%	31.6%	43.3%	35.5%	51.0%
Faith school	27.0%	23.4%	30.7%	39.0%	34.8%	43.3%
No disruptive peers in child's classroom	28.9%	26.6%	31.3%	42.4%	38.9%	45.8%
Disruptive peers in child's classroom	30.9%	27.4%	34.4%	41.9%	38.1%	45.6%
Number of classes in child's year						
Lowest tertile (Mean=1)	23.6%	20.6%	26.5%	31.6%	28.1%	35.0%
Medium tertile (Mean=2)	32.6%	29.4%	35.8%	48.7%	44.4%	53.1%
Highest tertile (Mean=3)	33.9%	29.2%	38.7%	47.8%	41.4%	54.3%
Number of children in child's class						

	Numeracy setting at MCS 4 (Age 7)			Numeracy setting at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Lowest tertile (Mean=20)	26.8%	23.8%	29.8%	38.2%	34.7%	41.7%
Medium tertile (Mean=27)	29.2%	26.2%	32.2%	46.3%	41.8%	50.7%
Highest tertile (Mean=30)	33.6%	29.7%	37.6%	40.7%	34.5%	46.8%
Child in single-year group	29.7%	27.3%	32.1%	42.6%	39.1%	46.1%
Child in mixed-year group	29.1%	25.3%	33.0%	40.9%	36.9%	44.9%
Number of SEN children in child's class						
Lowest tertile (Mean=0)	27.7%	25.0%	30.4%	40.2%	36.4%	44.0%
Medium tertile (Mean=1)	31.1%	26.9%	35.3%	42.3%	38.6%	46.0%
Highest tertile (Mean=5)	31.6%	28.0%	35.2%	45.9%	41.5%	50.3%
Number of EAL children in child's class						
Lowest tertile (Mean=0)	26.6%	23.8%	29.4%	39.7%	36.2%	43.3%
Medium tertile (Mean=1)	29.8%	26.4%	33.3%	38.6%	34.0%	43.2%

	Numeracy setting at MCS 4 (Age 7)			Numeracy setting at MCS 5 (Age 11)		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
Highest tertile (Mean=8)	34.3%	30.5%	38.2%	49.7%	44.8%	54.7%
Years of teaching experience						
Lowest tertile (Mean=3.8 years)	26.8%	23.8%	29.8%	38.2%	34.7%	41.7%
Medium tertile (Mean=12.6 years)	29.2%	26.2%	32.2%	46.3%	41.8%	50.7%
Highest tertile (Mean=22.7 years)	33.6%	29.7%	37.6%	40.7%	34.5%	46.8%

4.3.1.3: Characteristics associated with attainment grouping in the English and Welsh primary school

4.3.1.3.1: Children who were streamed

On average, during MCS 4 and MCS 5, children were not split evenly between top, middle, and bottom streams with 9% being placed in the top stream, 6.9% in the middle stream, and 4.6% in the bottom stream (see Table 4.7).

A similar proportion of boys and girls were in middle stream (6.6% boys, 7.1% girls), but girls were slightly overrepresented in the top stream groups (8.4% boys, 9.5% girls) and boys were overrepresented in the bottom stream groups (5.5% boys, 3.6% girls). Table 4.8 shows that girls were 33% (RRR 0.67, 95% CI 0.54, 0.83) less likely to be in the bottom stream groups.

Among the children who were streamed, there was significant difference in the ethnic composition of children in the top, middle or bottom streams. Compared to White MCS participants, Bangladeshi children were more likely to be streamed, while Black Caribbean children were over-represented in the middle and bottom stream groups.

Autumn babies (September–November births) were overrepresented in the top stream (11.4%), compared with 5.2% of the children in middle and 3.6% of children in the bottom streams. Younger babies (June–August births) were over-represented in the middle and bottom streams.

A higher percentage of children who scored high at the British Ability Scale Naming vocabulary test at MCS 3 (age 5) were in the top stream, while children who scored lower were overrepresented in the middle and bottom stream groups. More children in the middle or bottom stream were rated by their parent to have higher externalising symptoms at MCS 4 and MCS 5.

Children who reported liking school a lot were overrepresented in top stream groups and underrepresented in the bottom stream groups. Children who were unhappy at school were overrepresented in the middle and bottom stream groups.

Finally, children in the top stream had more advantageous socio-economic family circumstances, children in the bottom stream more disadvantages. Particularly, children in the bottom stream groups were the most likely to be in the bottom of the family income distribution. Parents of children in the bottom stream were also most likely to have no qualifications and least likely to have a degree level or higher qualification (NVQ4 or 5 equivalent), while children in single-parent households were more likely to be in the bottom stream groups compared to two-parent households.

Table 4. 7: Pooled RRRs of allocation to top, middle, and bottom streams at MCS 4 and MCS 5 according to child and family characteristics.

	Top stream					Middle stream					Bottom stream				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Gender															
Boys (Ref.)	8.4%					6.6%					5.5%				
Girls	9.5%	1.16*	0.08	1.02	1.33	7.1%	1.11	0.09	0.95	1.30	3.6%	0.67***	0.07	0.54	0.83
Ethnic group															
White (Ref.)	8.7%					6.5%					4.3%				
Mixed	10.2%	1.24	0.23	0.85	1.79	7.3%	1.02	0.25	0.62	1.66	7.0%	1.48	0.38	0.90	2.44
Indian	9.1%	1.34	0.25	0.93	1.93	7.0%	1.25	0.28	0.80	1.94	6.0%	1.21	0.37	0.67	2.19
Pakistani	8.7%	1.08	0.19	0.77	1.52	10.8%	1.96***	0.33	1.41	2.72	5.4%	1.39	0.34	0.85	2.25
Bangladeshi	17.8%	2.84***	0.62	1.86	4.36	16.4%	2.83***	0.72	1.72	4.68	7.0%	2.36*	0.86	1.15	4.84
Black Caribbean	9.6%	0.90	0.39	0.38	2.11	11.0%	2.67**	0.78	1.50	4.74	7.8%	2.59*	0.97	1.24	5.42
Black African	10.0%	1.19	0.26	0.77	1.83	8.7%	1.50	0.36	0.94	2.39	5.3%	1.66	0.60	0.81	3.40
Other	10.4%	0.97	0.25	0.58	1.62	4.2%	0.53	0.23	0.23	1.25	4.6%	1.02	0.44	0.43	2.41
Time of the year born															
Autumn 2000 (Ref.)	11.4%					5.2%					3.4%				
Winter 2000/1	10.4%	0.92	0.08	0.770	1.091	6.6%	1.11	0.14	0.87	1.42	4.5%	1.45*	0.25	1.04	2.03
Spring 2001	7.3%	0.70***	0.06	0.586	0.835	7.7%	1.45**	0.17	1.16	1.82	4.7%	1.44*	0.24	1.04	1.99
Autumn 2001	6.7%	0.60***	0.06	0.490	0.725	7.9%	1.47**	0.16	1.18	1.82	5.8%	1.89	0.30	1.39	2.58
Socio-economic characteristics															
Family income															
Highest incomes (Ref.)	11.4%					4.9%					2.3%				
2nd	10.3%	0.99	0.09	0.83	1.17	5.7%	1.18	0.16	0.91	1.53	3.5%	1.60*	0.29	1.12	2.29

	Top stream					Middle stream					Bottom stream				
	%	RRR	S.E	95% CI	95% CI	%	RRR	S.E	95% CI	95% CI	%	RRR	S.E	95% CI	95% CI
3rd	8.9%	0.82*	0.08	0.68	1.00	7.2%	1.43**	0.18	1.12	1.83	4.2%	1.95***	0.36	1.35	2.82
4th	6.9%	0.69***	0.08	0.55	0.86	7.9%	1.56**	0.21	1.20	2.03	5.6%	2.60***	0.47	1.83	3.70
Lowest incomes	6.6%	0.65***	0.08	0.52	0.82	9.0%	1.93***	0.25	1.50	2.48	7.9%	3.87***	0.67	2.75	5.45
Highest parental qualifications															
NVQ 4/5 (Ref.)	10.8%					6.0%					2.5%				
NVQ 3	8.9%	0.87	0.09	0.71	1.06	6.4%	1.05	0.14	0.81	1.37	4.3%	1.57*	0.28	1.10	2.23
NVQ 2	8.0%	0.77**	0.06	0.66	0.91	7.1%	1.16	0.13	0.94	1.43	4.3%	1.88***	0.28	1.41	2.51
NVQ 1	8.8%	0.89	0.12	0.67	1.17	8.3%	1.38*	0.22	1.01	1.88	7.7%	2.75***	0.55	1.86	4.08
Overseas qualifications	9.0%	0.71	0.15	0.47	1.07	8.2%	1.29	0.28	0.84	1.97	7.3%	2.98***	0.78	1.79	4.98
None	5.4%	0.61**	0.10	0.45	0.83	7.9%	1.45**	0.21	1.10	1.92	9.2%	4.00***	0.69	2.85	5.61
Family structure															
Two-parent family (Ref.)	9.5%					6.8%					4.0%				
Single-parent family	7.3%	0.76**	0.07	0.65	0.91	7.1%	1.00	0.10	0.81	1.22	6.3%	1.63***	0.18	1.31	2.04
Verbal skills at MCS 3 (quintiles)															
Highest (Ref.)	12.3%					3.8%					1.8%				
2nd	10.8%	0.94	0.09	0.77	1.14	6.0%	1.63**	0.25	1.20	2.21	2.3%	1.43	0.35	0.88	2.33
3rd	10.0%	0.91	0.10	0.74	1.12	6.5%	1.66**	0.27	1.21	2.28	2.8%	1.68*	0.42	1.04	2.74

	Top stream					Middle stream					Bottom stream				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
4th	7.8%	0.67***	0.07	0.54	0.82	7.5%	2.06***	0.30	1.54	2.75	4.8%	2.98***	0.67	1.92	4.62
Lowest	5.5%	0.55***	0.06	0.44	0.68	9.5%	2.85***	0.40	2.16	3.75	8.2%	5.68***	1.22	3.72	8.68
Parent reported externalising symptoms (quintiles)															
Lowest (Ref.)	12.7%					4.3%					1.2%				
2nd	10.4%	0.80*	0.07	0.68	0.95	7.3%	1.54**	0.19	1.21	1.97	3.3%	2.40***	0.55	1.53	3.76
3rd	8.7%	0.73**	0.07	0.60	0.87	6.9%	1.46**	0.19	1.13	1.89	3.5%	2.82***	0.66	1.79	4.45
4th	7.3%	0.54***	0.06	0.43	0.67	7.2%	1.68***	0.24	1.27	2.21	4.4%	3.29***	0.75	2.10	5.16
Highest	5.5%	0.41***	0.05	0.33	0.52	7.6%	1.77***	0.25	1.35	2.33	9.3%	7.03***	1.56	4.54	10.87
School experiences															
Unhappy at school															
All the time (Reference)	2.1%					7.5%					8.5%				
Some/most of time	8.7%	3.74***	1.35	1.83	7.63	6.8%	0.93	0.19	0.61	1.40	4.6%	0.47***	0.09	0.33	0.68
Never	10.3%	4.38***	1.59	2.15	8.93	6.8%	0.96	0.21	0.62	1.48	4.3%	0.44***	0.09	0.29	0.65
Tries best at school															
All the time (Ref.)	9.5%					6.6%					4.4%				
Some/most of time	8.1%	0.83*	0.06	0.71	0.95	7.4%	1.11	0.09	0.94	1.31	5.0%	1.19	0.13	0.95	1.47

	Top stream					Middle stream					Bottom stream				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Never	2.5%	0.29	0.19	0.08	1.03	8.4%	1.43	0.45	0.77	2.67	7.1%	1.94	0.73	0.93	4.07
Child likes school															
A lot (Ref.)	10.4%					6.9%					3.7%				
A bit	8.3%	0.73***	0.05	0.64	0.84	6.8%	1.00	0.08	0.86	1.17	0.962	4.7%	1.21	0.13	0.98
Does not like it	4.4%	0.41***	0.06	0.31	0.56	6.8%	1.02	0.14	0.79	1.34	0.856	8.2%	2.09***	0.32	1.55

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

4.3.1.3.2: Children who were set for literacy or maths

The results for set placement for literacy or maths (Tables 4.8 and 4.9) were similar to the results for children being in a streamed class. There were however some notable differences.

Contrary to children who were streamed or set for literacy, there were no statistically significant differences between boys and girls regarding allocation to the top, middle or bottom literacy sets.

In contrast to streaming, Indian boys and girls as well as boys and girls from the other ethnic group including Chinese students were overrepresented in the top set groups for literacy and numeracy.

Table 4. 8: Pooled RRRs for allocation to top, middle, and bottom literacy sets at MCS4 and MCS 5 according to child and family characteristics.

	Top literacy group					**Middle literacy group**					**Bottom literacy group**				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Gender															
Boys (Ref.)	13.6%					12.8%					11.5%				
Girls	18.1%	1.25***	0.07	1.12	1.39	11.9%	0.91	0.06	0.81	1.03	6.6%	0.56***	0.04	0.48	0.66
Ethnic group															
White (Ref.)	15.4%					11.7%					8.8%				
Mixed	18.5%	1.32	0.19	0.99	1.75	14.2%	1.34	0.23	0.95	1.89	8.4%	1.14	0.28	0.71	1.83
Indian	16.7%	1.50*	0.24	1.10	2.05	13.9%	1.55*	0.28	1.09	2.22	10.1%	1.27	0.30	0.80	2.02
Pakistani	17.4%	1.57**	0.21	1.20	2.05	15.8%	2.03**	0.30	1.52	2.70	13.8%	2.14***	0.36	1.53	2.99
Bangladeshi	23.8%	3.08***	0.68	1.99	4.76	27.0%	4.68***	0.99	3.09	7.11	13.4%	2.85**	0.87	1.57	5.18
Black Caribbean	16.8%	1.31	0.40	0.72	2.40	11.8%	1.74	0.54	0.95	3.20	11.4%	2.55**	0.78	1.40	4.64
Black African	15.4%	1.26	0.25	0.85	1.87	19.4%	1.93**	0.38	1.31	2.85	9.9%	1.66	0.46	0.97	2.86
Other	21.2%	1.56*	0.28	1.10	2.23	12.7%	1.05	0.24	0.68	1.64	8.8%	1.18	0.37	0.64	2.18
Time of the year born															
Autumn 2000 (Ref.)	19.7%					10.8%					6.3%				
Winter 2000/1	17.8%	0.93	0.07	0.80	1.07	11.7%	0.99	0.09	0.82	1.19	7.7%	1.28*	0.15	1.02	1.62
Spring 2001	13.5%	0.77**	0.06	0.66	0.90	13.5%	1.24*	0.11	1.04	1.48	9.3%	1.48**	0.17	1.17	1.87

	Top literacy group					Middle literacy group					Bottom literacy group				
	%	RRR	S.E	95% CI	95% CI	%	RRR	S.E	95% CI	95% CI	%	RRR	S.E	95% CI	95% CI
Autumn 2001	12.3%	0.64***	0.05	0.55	0.75	13.3%	1.20*	0.11	1.00	1.44	12.9%	2.10***	0.24	1.68	2.63
Socio-economic characteristics															
Family income															
Highest incomes (Ref.)	21.3%					10.0%					4.6%				
2nd	16.7%	0.89	0.07	0.77	1.03	11.2%	1.19	0.11	0.99	1.43	7.1%	1.66***	0.23	1.26	2.18
3rd	14.4%	0.72***	0.06	0.61	0.84	13.6%	1.38**	0.14	1.14	1.68	8.1%	1.95***	0.26	1.50	2.53
4th	14.4%	0.83*	0.07	0.70	0.98	12.8%	1.50***	0.15	1.24	1.83	12.5%	3.13***	0.41	2.42	4.04
Lowest incomes	11.2%	0.70***	0.06	0.59	0.82	14.6%	1.85***	0.18	1.52	2.25	14.2%	3.90***	0.51	3.02	5.04
Highest parental qualifications															
NVQ 4/5 (Ref.)	19.3%					10.8%					6.4%				
NVQ 3	14.6%	0.78**	0.06	0.66	0.91	12.1%	1.12	0.10	0.94	1.34	7.3%	1.13	0.14	0.88	1.44
NVQ 2	14.7%	0.80**	0.06	0.70	0.92	13.3%	1.26**	0.10	1.08	1.48	9.3%	1.53***	0.18	1.22	1.92
NVQ 1	12.1%	0.74*	0.09	0.58	0.94	14.0%	1.41*	0.19	1.08	1.83	13.6%	2.13***	0.31	1.59	2.84

	Top literacy group					Middle literacy group					Bottom literacy group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Overseas qualifications	14.0%	0.82	0.12	0.61	1.10	12.4%	1.51**	0.23	1.12	2.03	11.6%	1.98***	0.37	1.37	2.85
None	11.6%	0.73**	0.08	0.60	0.90	14.1%	1.46*	0.17	1.17	1.83	15.6%	2.83***	0.36	2.21	3.62
Family structure															
Two-parent family (Ref.)	16.4%					12.1%					8.3%				
Single-parent family	13.8%	0.90	0.06	0.78	1.03	12.9%	1.10	0.08	0.95	1.28	11.5%	1.54***	0.14	1.28	1.84
Verbal skills at MCS 3 (quintiles)															
Highest (Ref.)	21.6%					8.4%					3.9%				
2nd	20.0%	1.01	0.08	0.87	1.18	12.2%	1.49***	0.16	1.21	1.84	4.3%	1.24	0.21	0.89	1.73
3rd	17.3%	0.87	0.07	0.73	1.02	10.8%	1.32*	0.15	1.05	1.64	6.9%	1.78*	0.30	1.28	2.47
4th	13.2%	0.70***	0.06	0.59	0.83	14.3%	1.83***	0.19	1.49	2.25	9.6%	2.61***	0.40	1.94	3.51
Lowest	10.0%	0.60***	0.05	0.51	0.72	14.5%	2.07***	0.22	1.69	2.54	15.9%	5.02***	0.73	3.77	6.69
Parent reported externalising symptoms (quintiles)															
Lowest (Ref.)	24.5%					9.3%					2.6%				
2nd	18.0%	0.78***	0.05	0.68	0.89	12.1%	1.33*	0.12	1.11	1.59	5.8%	2.16***	0.34	1.58	2.95
3rd	15.0%	0.70***	0.05	0.60	0.81	13.0%	1.41***	0.13	1.17	1.70	8.3%	3.32***	0.54	2.42	4.56
4th	13.4%	0.60***	0.05	0.51	0.71	13.7%	1.63***	0.17	1.33	2.00	10.4%	4.19***	0.68	3.05	5.76
Highest	8.7%	0.40***	0.04	0.34	0.48	13.8%	1.66***	0.17	1.36	2.03	16.5%	6.74***	1.06	4.95	9.18

	Top literacy group					Middle literacy group					Bottom literacy group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
School experiences															
Unhappy at school															
All the time (Ref.)	4.6%					11.2%					16.3%				
Some/most of time	16.2%	3.59***	0.88	2.21	5.83	12.5%	1.22	0.20	0.88	1.68	9.4%	0.57***	0.08	0.43	0.76
Never	16.3%	3.47***	0.87	2.12	5.67	12.1%	1.20	0.21	0.85	1.70	7.8%	0.46***	0.07	0.33	0.63
Tries best at school															
All the time (Ref.)	16.3%					11.7%					8.4%				
Some/most of time	15.2%	0.98	0.05	0.87	1.09	13.6%	1.21**	0.08	1.07	1.37	10.3%	1.33***	0.10	1.14	1.54
Never	4.1%	0.33**	0.13	0.15	0.72	14.6%	1.22	0.37	0.68	2.20	16.7%	2.19**	0.61	1.27	3.78
Child likes school															
A lot (Ref.)	18.0%					11.6%					7.0%				
A bit	14.8%	0.86**	0.05	0.77	0.95	13.1%	1.13*	0.07	1.01	1.28	10.1%	1.39***	0.11	1.19	1.63
Does not like it	9.0%	0.52***	0.06	0.42	0.65	12.8%	1.10	0.12	0.89	1.35	14.7%	2.15***	0.24	1.74	2.67

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table 4. 9: Pooled RRRs for allocation to top, middle, and bottom sets for maths at MCS 4 and MCS 5 according to child and family characteristics.

	Top maths group					Middle maths group					Bottom maths group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Gender															
Boys (Ref)	22.6%					15.5%					12.4%				
Girls	22.2%	0.99	0.06	0.88	1.12	17.3%	1.12	0.08	0.97	1.27	11.4%	0.93	0.08	0.78	1.11
Ethnic group															
White (Ref.)	22.2%					15.9%					11.7%				
Mixed	24.0%	1.29	0.17	0.99	1.67	17.0%	1.26	0.20	0.92	1.72	12.7%	1.13	0.23	0.76	1.68
Indian	27.6%	2.02**	0.29	1.52	2.67	19.7%	1.81**	0.32	1.27	2.57	13.2%	1.67*	0.37	1.08	2.59
Pakistani	20.3%	1.18	0.16	0.89	1.55	18.8%	1.57***	0.23	1.17	2.11	13.7%	1.44*	0.23	1.06	1.97
Bangladeshi	25.2%	2.09***	0.40	1.44	3.04	27.8%	2.95***	0.65	1.91	4.54	14.9%	2.12**	0.56	1.26	3.57
Black Caribbean	19.7%	0.90	0.27	0.50	1.64	17.8%	1.63	0.44	0.96	2.77	17.4%	2.31**	0.68	1.30	4.13
Black African	21.3%	1.17	0.20	0.84	1.62	21.7%	1.59*	0.31	1.08	2.34	11.4%	1.21	0.31	0.74	2.00
Other	30.0%	1.70**	0.29	1.22	2.36	18.3%	1.21	0.28	0.77	1.89	7.8%	0.86	0.26	0.47	1.56
Time of the year born															
Autumn 2000 (Ref.)	28.9%					14.0%					7.7%				
Winter 2000/1	24.9%	0.89	0.06	0.78	1.01	14.8%	1.05	0.09	0.89	1.24	11.4%	1.36**	0.14	1.11	1.67
Spring 2001	18.9%	0.72***	0.05	0.63	0.82	17.8%	1.30**	0.10	1.11	1.52	13.0%	1.61***	0.17	1.31	1.98

	Top maths group					Middle maths group					Bottom maths group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Autumn 2001	17.2%	0.64***	0.05	0.56	0.74	18.9%	1.33***	0.11	1.14	1.55	15.5%	2.02***	0.21	1.65	2.48
Socio-economic characteristics															
Family income															
Highest incomes (Ref.)	8.2%					14.8%					8.2%				
2nd	9.9%	0.89	0.06	0.78	1.02	16.4%	1.13	0.09	0.96	1.33	9.9%	1.30*	0.15	1.04	1.64
3rd	11.8%	0.70***	0.05	0.61	0.80	15.9%	1.04	0.09	0.88	1.23	11.8%	1.41**	0.16	1.13	1.76
4th	14.4%	0.60***	0.05	0.52	0.71	16.6%	1.14	0.11	0.95	1.37	14.4%	1.70***	0.18	1.38	2.11
Lowest incomes	16.1%	0.54***	0.04	0.46	0.64	18.5%	1.33**	0.13	1.11	1.61	16.1%	2.14***	0.23	1.73	2.64
Highest parental qualifications															
NVQ 4/5 (Ref.)	28.4%					15.1%					9.1%				
NVQ 3	21.8%	0.74***	0.05	0.65	0.86	14.7%	0.92	0.08	0.77	1.11	11.1%	1.10	0.12	0.89	1.38
NVQ 2	20.5%	0.71***	0.05	0.63	0.81	18.3%	1.24**	0.09	1.06	1.43	11.5%	1.32**	0.12	1.10	1.59
NVQ 1	17.7%	0.66***	0.07	0.53	0.82	17.4%	1.18	0.14	0.93	1.49	16.7%	1.87***	0.23	1.46	2.38

	Top maths group					Middle maths group					Bottom maths group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Overseas qualifications	21.0%	0.80	0.11	0.61	1.04	14.5%	1.22	0.19	0.91	1.65	14.5%	1.76***	0.28	1.29	2.41
None	12.6%	0.50***	0.05	0.41	0.61	18.0%	1.22	0.13	0.99	1.49	18.5%	2.17***	0.23	1.76	2.66
Family structure															
Two-parent family (Ref.)	19.0%					16.9%					11.0%				
Single-parent family	16.3%	0.85**	0.05	0.75	0.96	16.8%	1.03	0.07	0.89	1.18	14.6%	1.32**	0.11	1.13	1.55
Verbal skills at MCS 3 (quintiles)															
Highest (Ref.)	30.4%					13.2%					5.7%				
2nd	28.9%	0.98	0.07	0.85	1.13	16.4%	1.29**	0.12	1.07	1.55	6.9%	1.41*	0.20	1.07	1.85
3rd	24.0%	0.81**	0.06	0.70	0.95	16.2%	1.15	0.11	0.95	1.39	9.4%	1.64***	0.23	1.25	2.16
4th	18.7%	0.66***	0.05	0.57	0.76	18.2%	1.36**	0.12	1.14	1.63	12.4%	2.30***	0.30	1.79	2.97
Lowest	14.4%	0.55***	0.04	0.47	0.65	17.6%	1.54***	0.14	1.29	1.83	20.5%	3.87***	0.48	3.03	4.95

	Top maths group					Middle maths group					Bottom maths group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Parent reported externalising symptoms (quintiles)															
Lowest (Ref.)	33.3%					12.9%					4.8%				
2nd	25.6%	0.82**	0.05	0.73	0.93	17.3%	1.39***	0.12	1.18	1.65	9.1%	1.90***	0.23	1.51	2.40
3rd	23.0%	0.75***	0.05	0.65	0.85	17.4%	1.44***	0.13	1.21	1.71	11.0%	2.58***	0.32	2.03	3.28
4th	18.7%	0.62***	0.05	0.53	0.73	17.8%	1.49***	0.15	1.23	1.81	14.5%	3.24***	0.40	2.55	4.12
Highest	13.1%	0.41***	0.04	0.35	0.49	17.0%	1.40***	0.13	1.16	1.68	20.1%	4.28***	0.50	3.40	5.37
School experiences															
Unhappy at school															
All the time (Ref.)	6.1%					13.5%					18.1%				
Some/most of time	23.1%	4.06***	0.80	2.76	5.97	16.9%	1.67**	0.29	1.19	2.34	12.6%	4.06***	0.80	2.76	5.97
Never	23.3%	3.94***	0.80	2.64	5.87	15.9%	1.49*	0.27	1.03	2.14	10.0%	3.94***	0.80	2.64	5.87

	Top maths group					Middle maths group					Bottom maths group				
	%	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**	**%**	**RRR**	**S.E**	**95% CI**	**95% CI**
Tries best at school															
All the time (Ref.)	22.1%					15.7%					10.9%				
Some/most of time	24.0%	1.22***	0.06	1.10	1.34	17.9%	1.29***	0.08	1.15	1.45	13.9%	1.48***	0.10	1.29	1.70
Never	7.9%	0.38**	0.12	0.20	0.70	17.2%	0.95	0.26	0.56	1.62	17.1%	1.63*	0.43	0.97	2.75
Child likes school															
A lot (Ref.)	24.5%					15.9%					9.8%				
A bit	22.4%	0.96	0.05	0.87	1.05	17.3%	1.14*	0.07	1.02	1.28	13.4%	1.40***	0.10	1.22	1.60
Does not like it	13.4%	0.50***	0.05	0.41	0.61	15.5%	0.91	0.09	0.75	1.10	16.3%	1.51***	0.16	1.23	1.86

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

4.3.1.4: The relationship between attainment grouping types in the MCS

There was a strong relationship between streaming and setting for literacy in schools. Figures 4.2 and 4.3 show that 67% of children who were streamed were also set for literacy at MCS 4 and 68% at MCS 5. Among the children not streamed, 21% were set for literacy at MCS 4 and 33% at MCS 5.

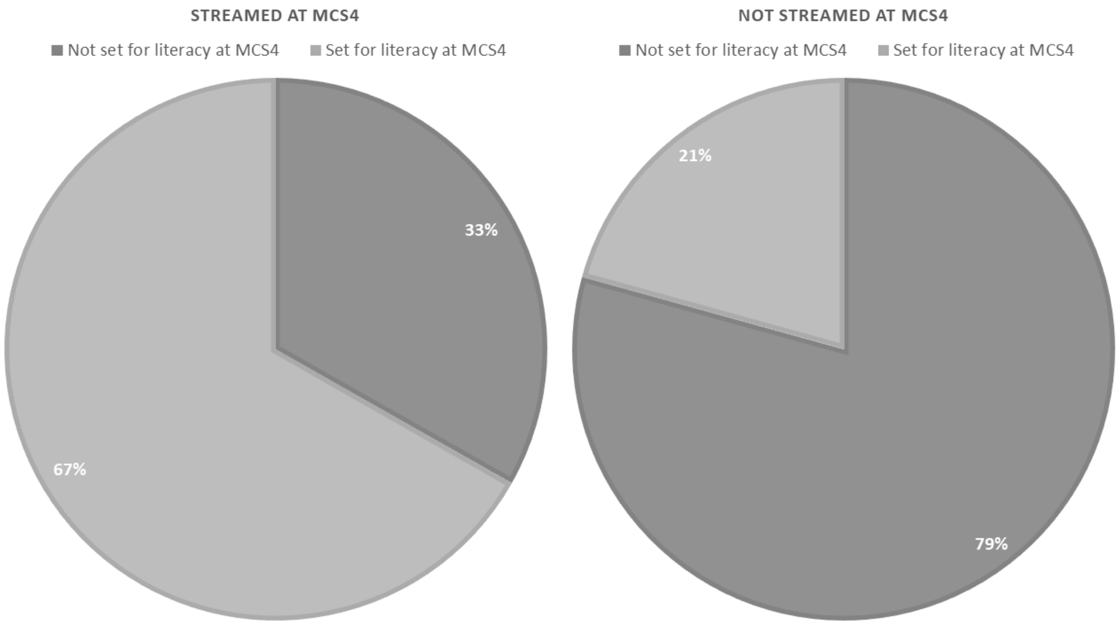

Figure 4. 2: Relationship between streaming and literacy setting at MCS 4

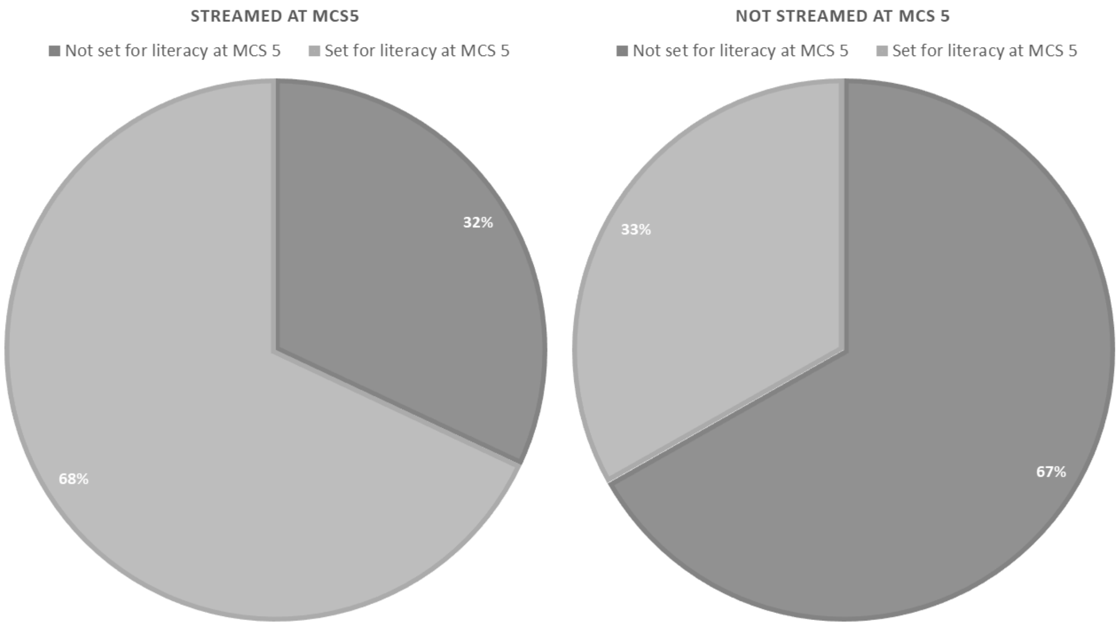

Figure 4. 3: Relationship between streaming and literacy setting at MCS 5

A strong relationship between streaming and setting for maths was also observed. Figures 4.4 and 4.5 highlight that 75% of children who were streamed were also set for maths at MCS 4 and 91% at MCS 5. Of those children not streamed, 26% were set for maths at MCS 4 and 55% at MCS 5.

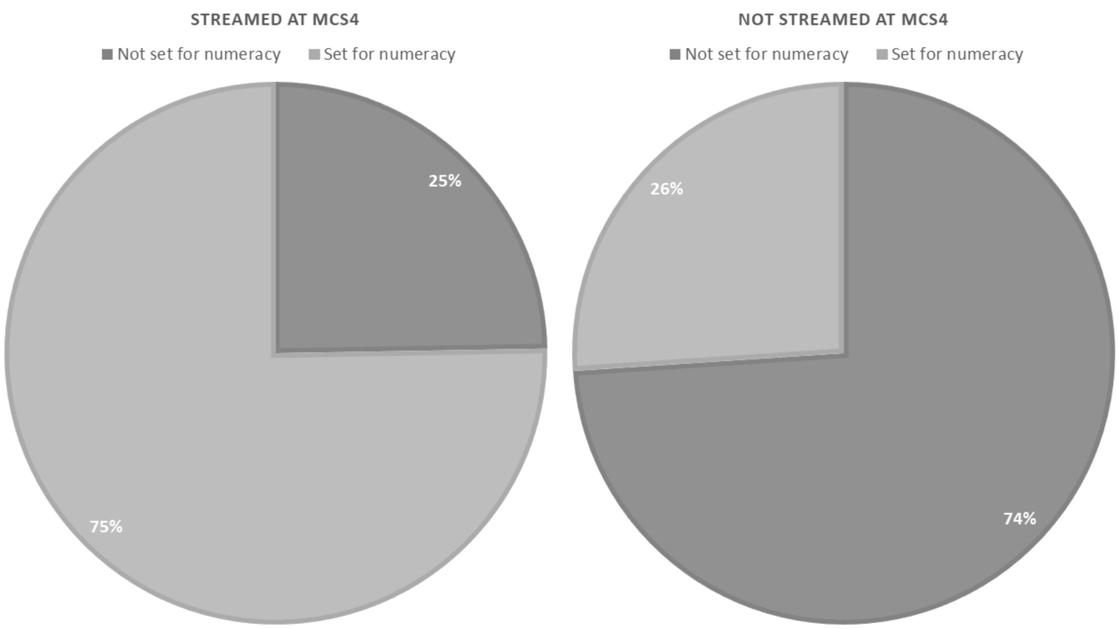

Figure 4. 4: Relationship between streaming and setting for maths at MCS 4

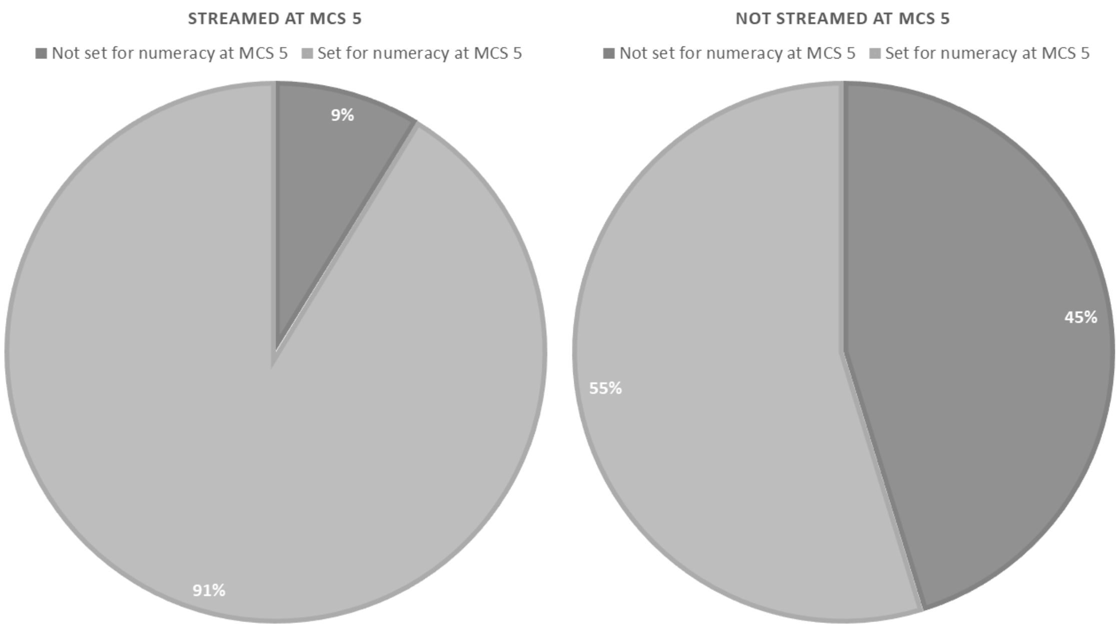

Figure 4. 5: Relationship between streaming and setting for maths at MCS 5

A strong relationship between setting for literacy and setting for maths was evident. Figures 4.6 and 4.7 show that 76% of children who were set for literacy were also set for maths at MCS 4 and 99% at MCS 5. Among children not set for literacy, 4% were set for maths at MCS 4 and 39% at MCS 5.

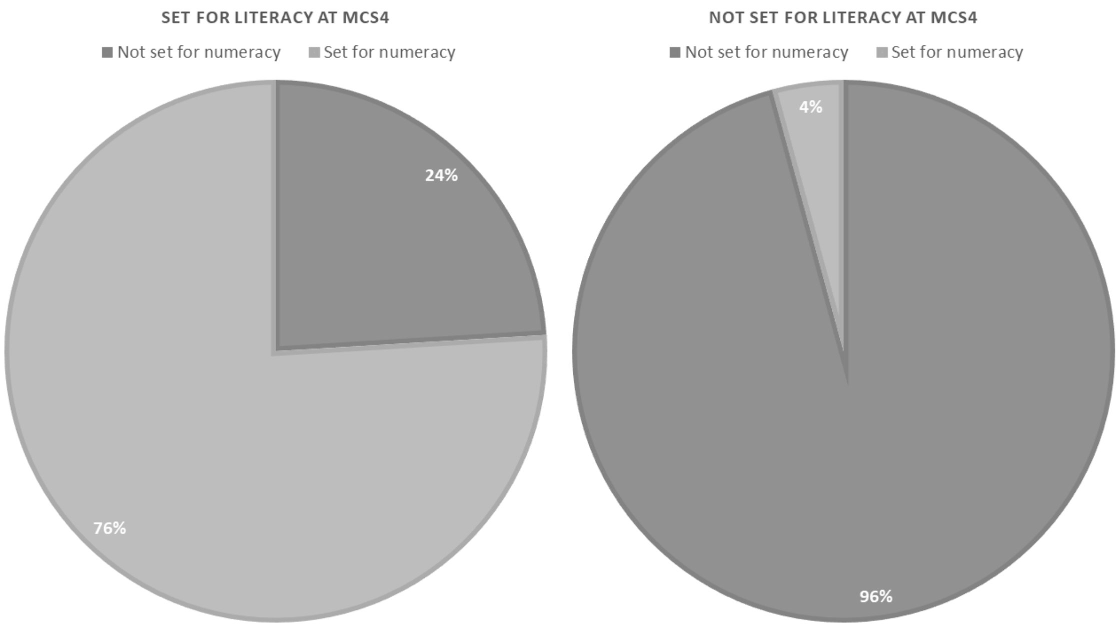

Figure 4. 6: Relationship between stetting for literacy and setting for maths at MCS 4

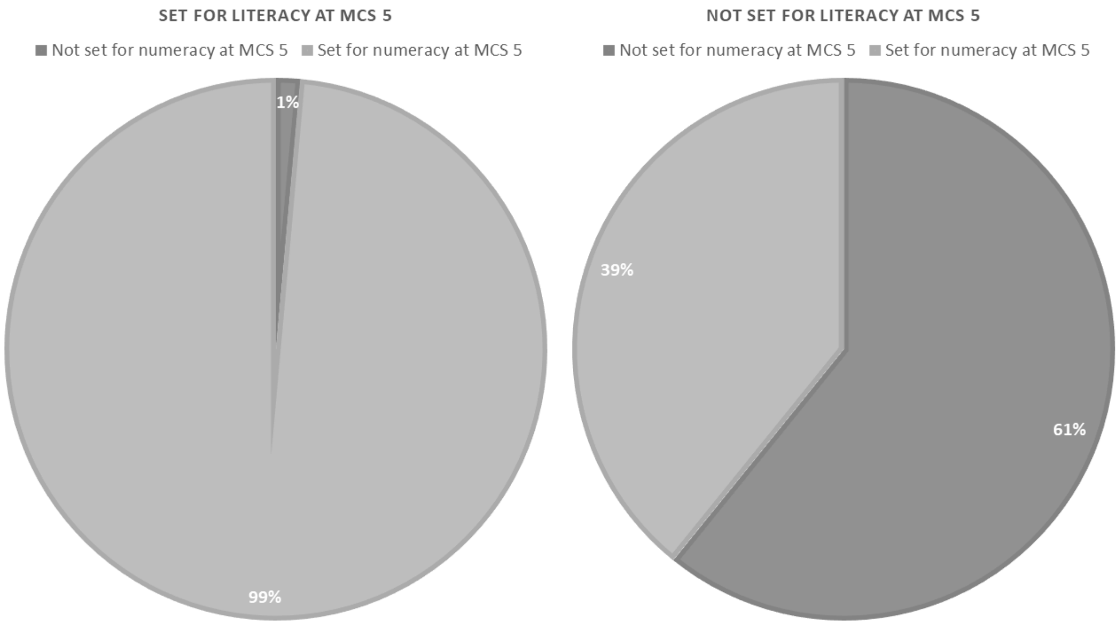

Figure 4. 7: Relationship between setting for literacy and setting for maths at MCS 5

Figure 4.8 shows a substantial overlap between streaming and setting for literacy and maths at both MCS 4 and MCS 5. As it is evident from Figure 4.8 very few participants were attending primary schools in England and Wales which were exclusively streaming or setting. In fact, most English and Welsh primary schools in the MCS practiced streaming and setting for literacy or/and maths in tandem.

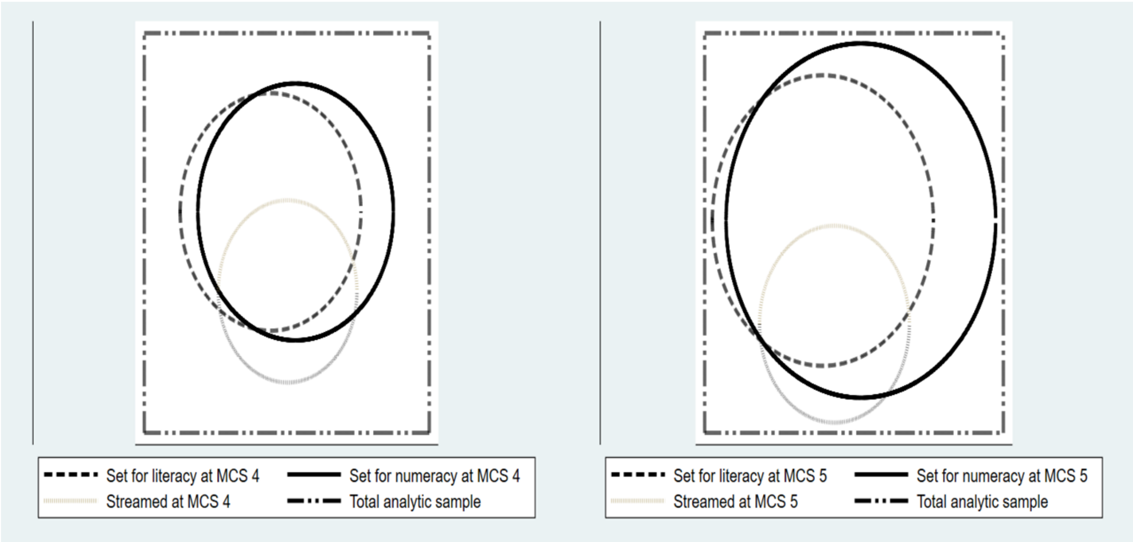

Figure 4. 8: Proportional Venn diagram of streaming and setting in English and Welsh primary schools at MCS 4 and MCS 5

4.3.1.4.1: Relationship between types of attainment grouping and group placement

Figures 4.9 to 4.11 show that the majority of children who were placed in the top or the bottom group of a specific attainment grouping type at MCS 4 and MCS 5 were also more likely to be placed in the top or the bottom group of another attainment group type. There was a stronger relationship between set placement for maths and stream, especially at MCS 5 (age 11), than set placement for literacy and stream. A weaker association was observed for set placement for maths and set placement for literacy (see Figure 4.11).

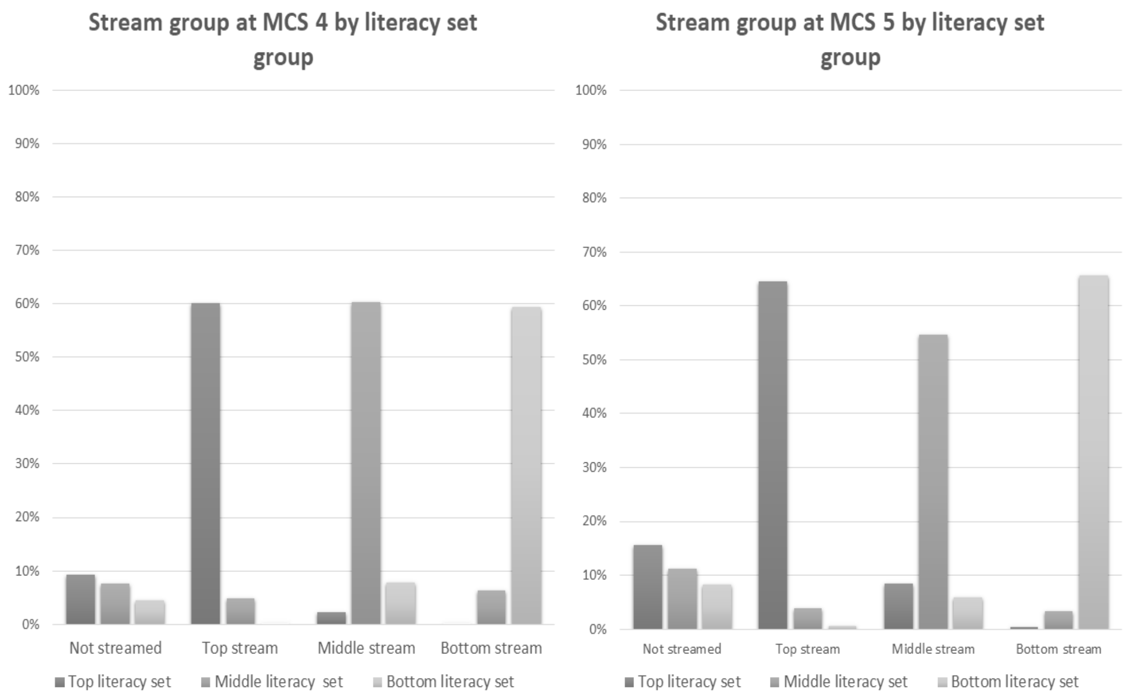

Figure 4. 9: Relationship between stream and literacy set groups at MCS 4 and MCS 5

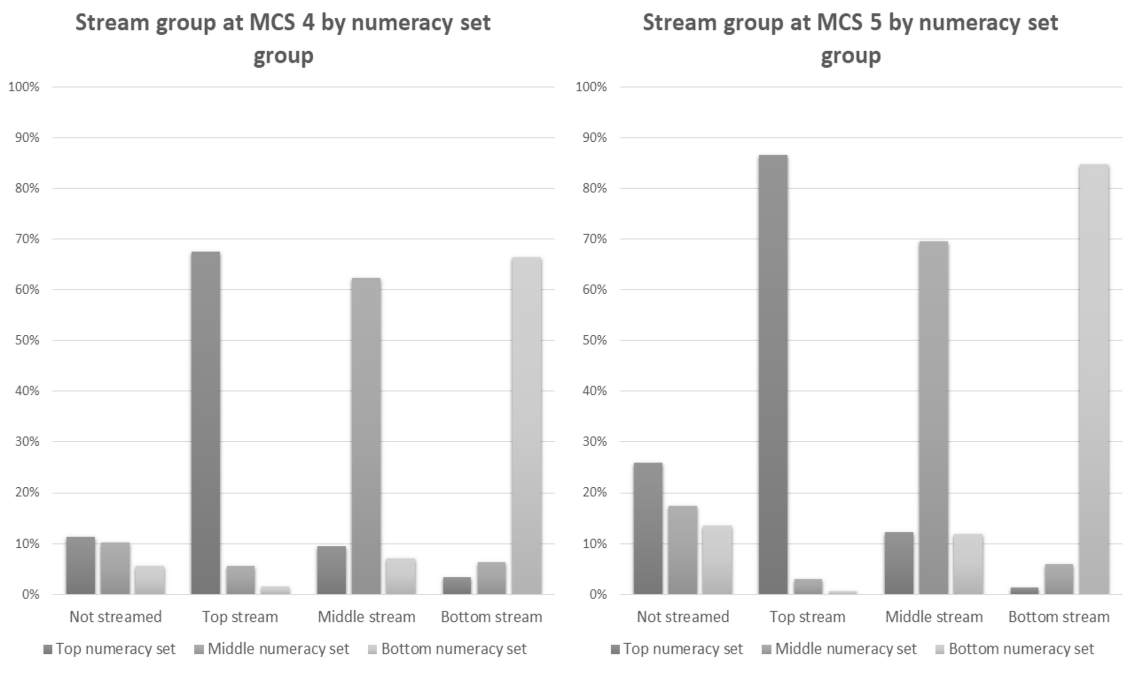

Figure 4. 10: Relationship between stream and maths set groups at MCS 4 and MCS 5

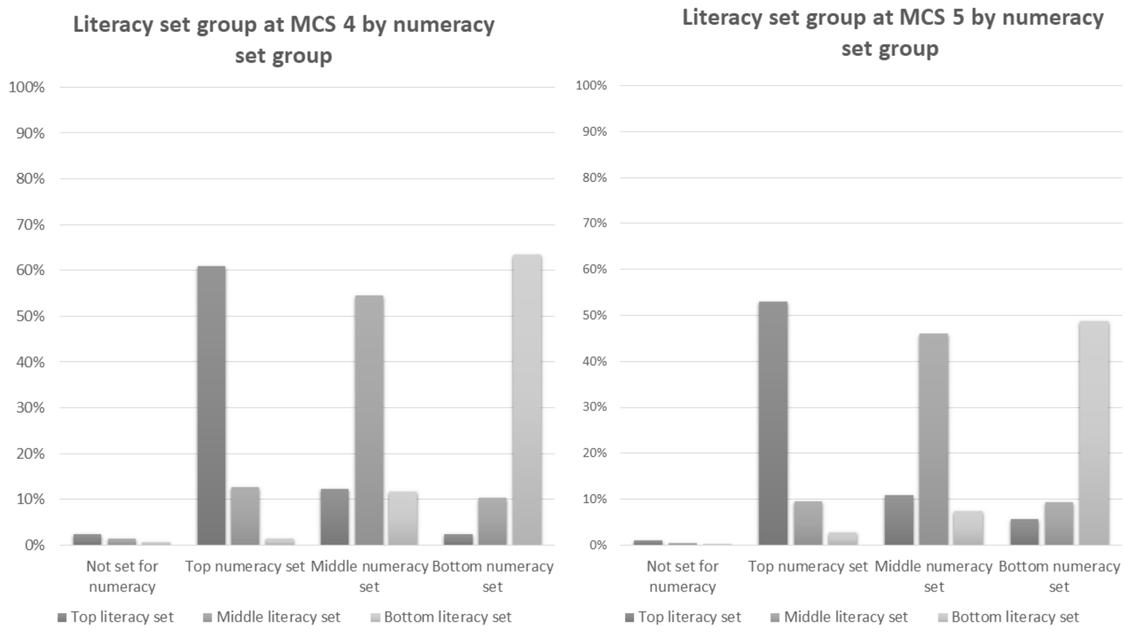

Figure 4. 11: Relationship between literacy set and maths set groups at MCS 4 and MCS 5

4.3.1.5: Attainment grouping transitions

The findings of the above analyses based on the MCS 4 and MCS 5, provide evidence that there is a substantial overlap between streaming and setting in English and Welsh primary schools. Therefore, the complexity of attainment grouping structures in English and Welsh primary schools poses a challenge in separating out the effects of different aspects of attainment grouping. For this reason, throughout the analyses of chapters 5 to 7 the term attainment grouping refers to the combined effects of streaming or setting for literacy/maths. In addition, streaming and setting are usually combined in the evidence reviews on the effect of attainment grouping (Education Endowment Foundation, 2018a).

A binary attainment grouping variable (yes/no) rather than a nominal categorical variable flagging placement in the top, middle or bottom streams or/and sets was favoured since this chapter's analyses highlighted that assigning students to different attainment groups is biased by children's social backgrounds and their individual characteristics. In an observational study as the MCS this could be an issue as there is always possibility of omitted teacher, child and family level confounding which could bias the associations of interest.

A substantial number of MCS participants who were not grouped by attainment at age 7 (MCS 4), were grouped by attainment at age 11 (MCS 5). This is important because 1) a transition to attainment grouping between age 7 and age 11 is likely to be a school-level decision and therefore not subject to observed and unobserved teacher, child and family level confounding, and 2) because a transition from mixed attainment to attainment grouping variable provides a baseline where the two groups of children were taught in mixed attainment groups at age 7. This allows effectively for comparisons between those children who were taught in mixed attainment groups at both age 7 and age 11, and those children who were taught in mixed attainment groups only at age 7, and change in their developmental outcomes from age 7 to age 11.

MCS participants who were taught in mixed attainment groups at age 7, but they were reported to be grouped by attainment at age 11 were assigned a value of 1. The comparison group included those MCS participants whose teachers reported that they were taught in mixed attainment groups at age 7and age 11 and were assigned a value of 0. MCS participants who were continuously grouped by attainment at both age 7 and age 11 and those who transitioned from attainment grouping to mixed grouping were not included in the statistical analyses of chapters 5 and 6.

Figure 4.12 demonstrates that only 21.1% (n=1,518) of MCS participants were not grouped by attainment (streaming and setting for literacy/numeracy) at ages 7 and 11 years, while 37.1% (n=2,670) were grouped only at age 11, 12.2% (n=878) were grouped by attainment only at age 7 and 29.6% (n=2,130) of MCS participants were grouped by attainment when they were both ages 7 and 11 years.

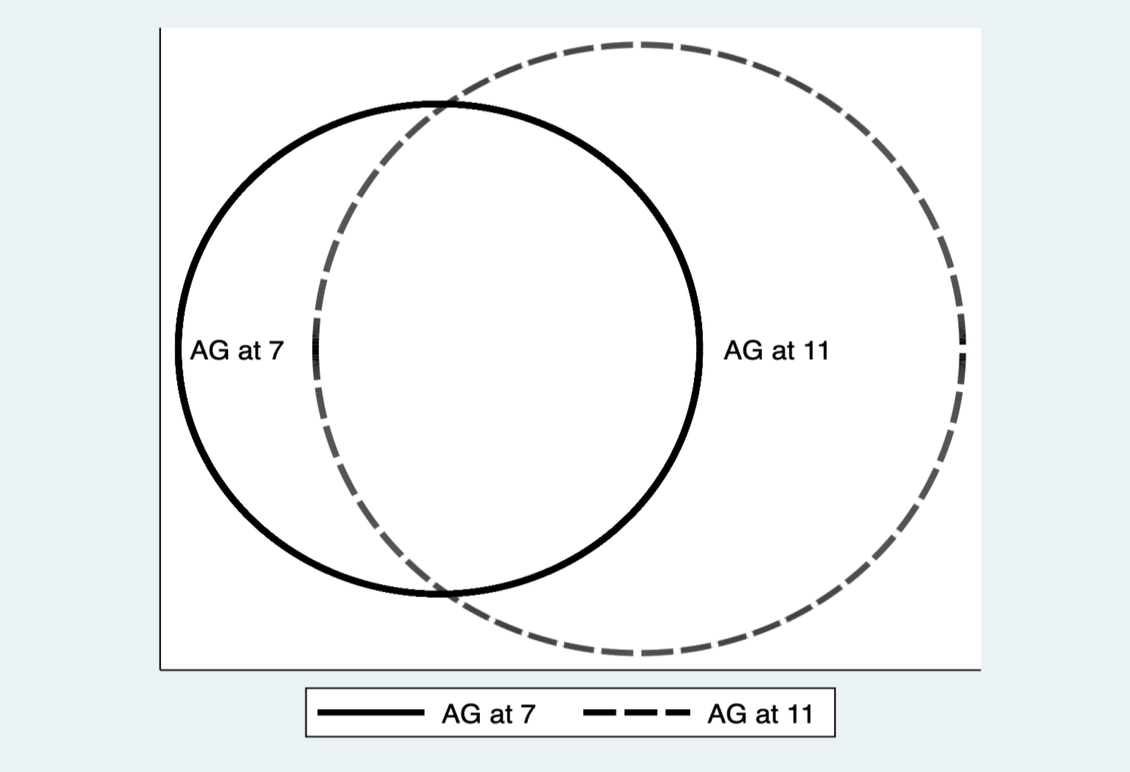

Figure 4. 12: Proportional Venn diagram of attainment grouping in English and Welsh primary schools at age 7 and 11 (N=7,196)

Figure 4.13 highlights a significantly greater proportion of children who changed to setting for maths between age 7 and age 11 compared to children who changed to setting for literacy or streaming. Additionally, a significantly higher proportion of children who changed from mixed attainment grouping at age 7 to attainment grouping at age 11 were placed at top streams or sets rather than middle or bottom streams or sets.

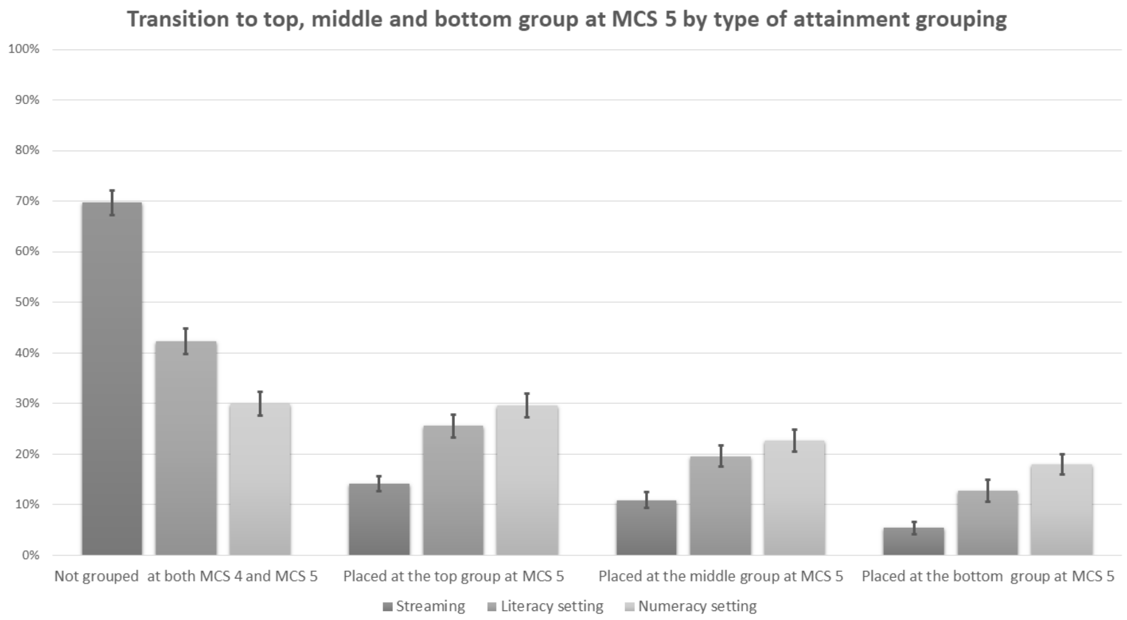

Figure 4. 13: Incidence of streaming and setting, and allocation to groups between MCS 4 and MCS 5

4.3.1.5.1: Attainment grouping transitions in the English and Welsh primary school by school related characteristics

Table 4.10 highlights that MCS participants in larger primary schools in England and Wales were more likely to change to attainment grouping between age 7 and age 11. Other characteristics related to larger primary school size such as larger class size, a higher number of children with EAL within child’s classroom, and mixed year classes indicated a higher probability of attainment grouping transition between age 7 and age 11.

Table 4. 10: Attainment grouping transitions between MCS 4 and MCS 5 by school-related characteristics

	Not streamed/set between MCS 4 and MCS 5			Streamed/set between MCS 4 and MCS 5			De-streamed/de-set between MCS 4 and MCS 5			Streamed/set at both MCS 4 and MCS 5		
	%	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**	**%**	**95% CI**	**95% CI**
School type												
State-maintained	19.8%	17.6%	22.0%	54.0%	51.4%	56.6%	6.8%	5.7%	8.0%	19.4%	17.0%	21.8%
Fee-paying	28.4%	20.9%	35.9%	48.8%	41.7%	56.0%	4.8%	2.3%	7.3%	17.9%	11.9%	24.0%
Faith school	24.7%	21.2%	28.2%	52.2%	48.1%	56.3%	8.7%	5.7%	11.6%	14.4%	11.9%	16.9%
Disruption in classroom from peers												
No disruptive peers in child's classroom	21.5%	19.3%	23.8%	56.4%	53.8%	59.0%	6.6%	5.3%	8.0%	15.5%	13.7%	17.2%
Disruptive peers in child's classroom	22.7%	20.0%	25.4%	54.8%	51.9%	57.7%	6.0%	5.0%	7.1%	16.5%	14.2%	18.8%
Number of classes in child's year												
Lowest tertile (Mean=1) (Ref.)	36.6%	33.2%	40.0%	42.9%	39.6%	46.1%	10.9%	9.1%	12.7%	9.6%	7.8%	11.4%
Medium tertile (Mean=2)	14.7%	12.5%	16.8%	55.2%	52.2%	58.3%	5.7%	4.5%	7.0%	24.4%	21.5%	27.3%
Highest tertile (Mean=3)	10.5%	7.2%	13.8%	55.4%	50.4%	60.4%	6.3%	2.1%	10.4%	27.8%	22.8%	32.7%

	Not streamed/set between MCS 4 and MCS 5			Streamed/set between MCS 4 and MCS 5			De-streamed/de-set between MCS 4 and MCS 5			Streamed/set at both MCS 4 and MCS 5		
Number of children in child's class												
Lowest tertile (Mean=20) (Ref.)	26.8%	24.0%	29.7%	47.5%	44.5%	50.6%	7.9%	0.07	0.09	17.8%	15.3%	20.2%
Medium tertile (Mean=27)	19.8%	16.9%	22.7%	52.4%	49.3%	55.6%	7.1%	0.06	0.09	20.6%	17.5%	23.8%
Highest tertile (Mean=30)	19.0%	16.3%	21.8%	52.2%	48.4%	56.0%	8.2%	0.05	0.11	20.5%	17.7%	23.3%
Number of SEN children in child's class												
Lowest tertile (Mean=0)	23.6%	21.0%	26.3%	49.7%	46.9%	52.6%	8.0%	6.3%	9.8%	18.6%	16.3%	20.9%
Medium tertile (Mean=1)	19.8%	17.2%	22.4%	51.4%	48.1%	54.7%	7.4%	5.7%	9.0%	21.5%	18.7%	24.2%
Highest tertile (Mean=5)	21.9%	18.9%	24.9%	50.8%	47.3%	54.3%	7.9%	6.3%	9.5%	19.4%	16.7%	22.2%
Number of EAL children in child's class												
Lowest tertile (Mean=0)	25.2%	22.5%	28.0%	48.3%	45.1%	51.5%	9.1%	6.7%	11.4%	17.4%	15.1%	19.6%
Medium tertile (Mean=1)	21.1%	18.3%	23.9%	50.1%	46.8%	53.4%	7.4%	6.0%	8.9%	21.4%	18.5%	24.3%

	Not streamed/set between MCS 4 and MCS 5			Streamed/set between MCS 4 and MCS 5			De-streamed/de-set between MCS 4 and MCS 5			Streamed/set at both MCS 4 and MCS 5		
Highest tertile (Mean=8)	16.7%	13.8%	19.6%	55.2%	51.7%	58.6%	6.1%	4.6%	7.5%	22.0%	19.0%	25.1%
Years of teaching experience												
Lowest tertile (Mean=3.8 years)	20.3%	17.6%	23.0%	49.9%	46.8%	53.0%	7.8%	6.3%	9.3%	22.0%	19.4%	24.6%
Medium tertile (Mean=12.6 years)	23.8%	20.8%	26.7%	50.7%	47.6%	53.8%	7.6%	6.1%	9.2%	17.9%	15.5%	20.4%
Highest tertile (Mean=22.7 years)	22.0%	19.4%	24.5%	51.7%	48.5%	54.8%	7.8%	6.1%	9.6%	18.5%	16.0%	21.1%
Child in single-year group	20.0%	18.0%	22.0%	57.9%	55.4%	60.3%	6.1%	4.9%	7.3%	16.0%	14.2%	17.8%
Child in mixed-year group	28.1%	24.9%	31.4%	49.3%	45.5%	53.1%	7.4%	5.8%	8.9%	15.2%	12.7%	17.7%

4.3.2: Verbal skills

Verbal ability was the only measure of cognition that was measured consistently from the age of 3 to age 11 years in the MCS. It was used for the analyses of chapter 5. The British Ability Scales 2nd Edition (BAS II) is a battery which could measure cognitive ability and educational achievement (Elliot et al., 1997). The BAS is a standardized assessment battery suitable for children and adolescents from 3 years to 18 years (Connely, 2013). It has high test-retest reliability (Elliot et al., 1997) and it is also considered to be compatible with current psychological practice (Hill, 2005). At age 7 years (MCS4) the word reading achievement scale, which assesses children's English reading ability, was utilised (Hansen et al., 2014). When the children were 11 years old (MCS5), the verbal similarities ability test was used to examine verbal reasoning and verbal knowledge (Gallop et al., 2013). T-scores or Standardized BAS scores for both item difficulty and age of child were generated using the BAS normative data set for each test, with higher scores indicating better language performance (Connely, 2013). T scores were converted from raw scores. These scores are simply the number of correct answers the child gives in each test. Their simplicity is misleading if their age and difficulty are not considered. Thus, they are converted to T-scores which account for child's age and item difficulty. A participant with a T score of 50 had the same raw score as the mean of those in their norming group. The BAS measures are nationally standardized and have good inter-rater reliability (Elliot et al., 1997). In this research BAS scores were used as continuous measures.

4.3.3: Psychological symptoms

In this research, MCS participants' psychological symptoms were measured through reports from the child's main carer and teacher (for a breakdown see table 4.5) using the SDQ , which is a brief behavioural screening questionnaire for children aged 3 to 17 years (Goodman, 1997). Screening tools like the SDQ are usually more inclusive than clinical diagnoses since children and adolescents fail to access or receive mental health care (Garratt et al., 2017). Both teacher and parent responses in SDQ, when MCS participants were around age 7 and age 11, were used for the analyses of chapter 6 for comparative and robustness purposes. The analyses of chapter 7 utilised only parent responses in SDQ since teacher SDQ was not administered at MCS6.

The SDQ consists of 25 items which are divided equally to 5 subscales; emotional symptoms; conduct difficulties; hyperactivity/inattention difficulties; peer relationship difficulties; and prosocial behaviour (Goodman, 2001). All item responses to the SDQ were scored as 0, not true; 1, somewhat true; or 2, certainly true. Emotional and peer subscales can be summed to formulate an internalising difficulties score, while conduct and hyperactivity subscales can be summed to formulate an externalising difficulties score, with a higher score indicating more internalising or externalising difficulties. The scores can lie between 0 (no difficulties at all) and 20 (respondent thinks that the child certainly has difficulties in all areas covered in the subscales of the questionnaire).

This research used both externalising and internalising symptoms scores as continuous measures. Externalising and internalising psychological symptoms are conceptually distinct, so they were examined separately as it might be possible to have specific associations with attainment grouping at age 7 or transition to attainment grouping between 7 and 11 years. Besides, it has been found that these two subscales have a greater validity than individual subscales when examining specific behaviours (Goodman et al., 2010).

Table 4. 11: The Strengths and Difficulties Questionnaire (SDQ)

Scale/category	**Attribute**
Emotional	Has many worries, often seems worried
	Often unhappy, downhearted, tearful
	Complains of headache/sickness
	Has many fears/is easily scared
	Nervous/Clingy in new situations
Peer problems	Picked or bullied by other children
	Often solitary, plays alone
	Has at least one good friend*
	Generally liked by other children*
	Gets on better with adults than children
Hyperactivity/inattention	Easily distracted, concentration wanders
	Seeks tasks through to end*
	Constantly fidgeting or squirming
	Restless, overactive, cannot stay still
	Thinks things through before acting*
Conduct	Often has temper tantrums
	Generally obedient*
	Fights with or bullies other children
	Often lies or cheats
	Steals from home, school, elsewhere
Prosocial	Considerate of other people's feelings
	Shares readily with other children
	Helpful if someone is hurt, upset or unwell
	Kind to younger children
	Often volunteers to help other children

*Reverse coded

4.5: Potential confounders

Individual and family-level confounding was not an issue for the statistical analyses of chapters 5 and 6 , since the decision of primary schools to adopt or not attainment grouping is not likely to be driven by characteristics of the child or his/her family as shown above, rather than raising attainment to meet government priorities, alongside managing practical issue in the school environment (Hallam et al., 2002, Hallam et al., 2004a). Confounding at the aggregate-level (school-level) was a matter of concern. Given that the analyses of chapters 5 and 6 focus on the difference in verbal skills changes of those children who were not grouped by attainment during KS 2 and those who were (parallel growths), time-invariant or school-related characteristics that did not change substantially between MCS 4 and MCS 5 (school size, classroom size, number of children with EAL in child's classroom ,and whether the child was taught in mixed-year classroom) were assumed not to have a considerable impact on the association between the transition to attainment grouping and the change in children's verbal skills and psychological symptoms, thus it was decided not to include them in the statistical models. However, time-invariant school characteristics could have affected substantially the statistical analyses of chapter 7 (association between attainment grouping at age 7 and trajectories of externalising and internalising symptoms from 7 to 14 years). MCS UK primary schools that adopted attainment grouping at age 7 might have been more likely to be compositionally different than those which did not adopt attainment grouping at age 7. As potential school-level confounders were considered available characteristics of MCS schools reported either from MCS participant's parent or teacher that did not vary much over time (e.g school type, school size etc.).

What follows is a description of potential parent-reported or teacher-reported time-invariant school-related confounders at age 7 which have been highlighted by previous research as important and which might account for selection to attainment grouping at age 7, baseline differences (age 7) and the rate of change in parent-reported externalising and internalising symptoms (from 7 to 14 years).

4.5.1: School type

Data from previous studies suggested that different types of schools choose to adopt attainment grouping based on meeting the needs of their student population. For instance, recent evidence from the MCS showed that state-maintained schools (community schools) in the UK were more likely to adopt streaming or setting in the early years of primary school compared to fee-paying primary schools (Hallam and Parsons, 2013b). It has also been demonstrated that children attending fee-paying and faith schools (they are funded by the state and the church) have better socio-emotional and cognitive outcomes compared to children attending non-faith state maintained (community) schools; however, there is increasing evidence that such effects may falsely be attributed to schools themselves, rather than to their pupil intake (Sullivan et al., 2018). Specifically, students attending fee-paying and faith schools may be doing better because they are exposed to more resourceful family and school environments, e.g. highly stimulating home environments, where parents devote time and vital resources onto child's optimal development or higher within-school resources devoted to students of fee-paying primary schools in the UK (Green and Kynaston, 2019).

4.5.2: Number of classes within child's school Year (School size)

Previous research has demonstrated that children attending larger primary schools are more likely to be streamed or set compared to children attending smaller schools (Hallam and Parsons, 2013a). The size of the school, and subsequently the availability of school resources, also seems to play a vital role in how children learn and behave (Gump, 1988, Steele et al., 2007). Larger schools in the UK (usually state-maintained schools) face multiple pressures from local authorities or external accountability bodies such as Ofsted in order to increase the attainment of specific student populations of interest (e.g. children from deprived communities), as a consequence they seem to be more prone in adopting attainment grouping. Also, previous studies have shown that children attending larger schools may be exposed to negative school cultures such as bullying with further possible ramifications on their social and emotional development (Bowes et al., 2009). One possible explanation is that large schools in the UK may have greater age ranges of pupils, increasing the risk for younger children being bullied by older pupils (Bowes et al., 2009). Since data on the actual size of an MCS child's school can only be accessed via secure arrangements made with the UK Data Service; a proxy variable of school size was used which was measured by asking teachers in MCS4 about the number of classes in child's school year. A continuous measure of this proxy variable for CM's school size at age 7 was used.

4.5.3: Teacher working experience

Teacher working experience may be an important factor in decisions made by schools regarding adopting attainment grouping. Research shows that more experienced teachers are less likely to move to a school located in more deprived areas or move to a school with a worse Ofsted grade (Department for Education, 2017). It is also likely that teachers in more deprived areas are also more likely to leave the profession for good (Department for Education, 2017). As a result more vacancies become available in larger schools in economically deprived areas which are more likely to be filled by younger and less experienced teachers (Allen et al., 2012). As a result, schools may adopt attainment grouping with the aim to match pupils' needs with their teachers' expertise to increase educational standards. The research to date has been able to convincingly show that teacher experience and skills may affect both the teacher's ability to assess student's ability and group children accordingly, and their abilities to assist learning and both socioemotional and cognitive outcomes of their students. Relatedly, evidence shows that teachers with many years of experience have better control of their students' behaviour (Ritter and Hancock, 2007). In this PhD, teacher working experience is measured by asking the cohort member's teacher about the number of years teaching altogether. A continuous measure of teacher working experience is then used in the regression analyses.

4.5.4: Classroom characteristics

A considerable body of evidence shows that children in mixed year group classes are more likely to be streamed or set as a strategy adopted by schools to meet children's needs and raise educational standards.

Associations between class size, learning and behaviour have been found in a number of studies (Blatchford et al., 2003, Finn et al., 2003, Blatchford et al., 2011), while others have found links between mixed year group classes and prosocial and aggressive forms of behaviour (Winsler et al., 2002). In the MCS, class size was reported by the child's teacher by asking "how many children are in child's class'' and it was used as a continuous measure in the regression analyses. Teachers were also asked to report if child's class contained a mixed year group or not.

4.5.5: Classroom peer effects

Classroom disruption peer effects may drive schools' decisions on adopting streaming or setting for certain subjects (Lazear, 2001). The definition of ''disruptive'' in this case is not limited to bad behaviour (Lavy et al., 2012). It includes a wider range of factors, such as the reallocation of time that a teacher has to teach in order to concentrate on specific learning needs of particular pupils, such as Special and Education Needs (SEN) or those of pupils with English as an Additional Language (EAL) needs with possible implications on aggregate learning and behaviour at the classroom. It is also likely that these variables reflect upon the effect of structural school characteristics related to school size. In the MCS, teachers were asked to report whether there were students with disruptive behaviour in cohort member's classroom (yes/no), the number of children in child's class with SEN and EAL. Both the number of pupils with SEN and EAL in an MCS child's classroom were used as continuous measures in the regression analyses.

4.6: Effect modifiers

The association between attainment grouping, verbal skills and externalising or internalising psychological symptoms might differ according to child's or family characteristics. Statistical analyses showed that effect of early attainment grouping (age 7) or the effect of the transition to attainment grouping (between ages 7 and 11 years) did not vary by gender. So, gender was used in all statistical analyses of this PhD as a predictor of verbal skills and psychological outcomes. What follows is a description and the rationale of using family income and initial verbal skills (age 5) as potential effect modifiers in chapters 5, 6 and 7.

4.6.1: Family income

To understand whether attainment grouping is associated with an increase of socio-economic inequalities; family income was used as an indicator of the socio-economic circumstances under which the child lives and develops. In the MCS, family income was calculated as the total family income from all sources after tax and before housing costs. It was then adjusted for family size and composition using the modified Organisation for Economic Cooperation and Development (OECD) scales for equivalisation. Each modified OECD scale for equivalisation sets the family needs relative to those of a couple with no children whose scale is equal to 1. For instance, a two-parent family and 2 children under 14 have a scale of 1.4. Equivalised family income was derived by the CLS by dividing the total family income by the equivalisation factor to best operationalize the family's material resources. Missing data on family income were also imputed by the CLS. Pre-derived categorical measures of family income (quintiles) were used. A categorical instead of a continuous measure of family income was preferred since there is evidence of a non-linear causal relationship between family income and child outcomes, supporting the theory that the causal effects of income on cognitive and behavioural outcomes are more pronounced among children from low-income families (Cooper and Stewart, 2013b, Cooper and Stewart, 2017).

4.6.2: Baseline verbal skills

Baseline verbal skills at MCS 3 (age 5 years) were assessed using the Naming Vocabulary subscale from the BAS. The BAS Naming Vocabulary assesses verbal ability/expressive language of children. Continuous raw scores were converted into T-scores or Standardized scores adjusting for both item difficulty and age. T-scores with a mean of 50 and a standard deviation of 10 were used.

4.7: Methods

4.7.1: Fixed Effects Difference-in-Difference panel models

Fixed-effects difference-in-difference modelling was applied to explore the relationship between transition to attainment grouping (between MCS 4 and MCS 5) and average changes in children's verbal and behavioural outcomes (between MCS 4 and MCS 5).

Linear fixed effects models were used because unlike binary or count models, they utilize the full range of cognitive and behavioural data and thus provide the most detailed examination of children's cognitive and behavioural outcomes (Mirowsky and Ross, 2002).

Difference- in-difference design has been extensively used in social sciences to assess the effectiveness of policies and programs in public health (Wharam et al., 2007, Branas et al., 2011, Galiani et al., 2005, Kelly et al., 2013a, McKinnon et al., 2015, Glymour et al., 2008), education (Jerrim and Sims, 2018, Hanushek and Wossmann, 2006, Duflo et al., 2011) and the economy (Card and Krueger, 1994). In the absence of Randomized Controlled Trials, researchers often seek answers from quasi-experimental methods such as 'difference-in-difference. This is because it is not typically feasible to learn about the effect of social exposures by randomly assigning them to a policy or policy change. Difference-in-difference design mimics a randomised control trial in an observational setting as closely as possible, but one without randomised assignment to treatment and the control groups (Kahn-Lang and Lang, 2018). Specifically, in a difference-in-difference design the only difference between those exposed and those not exposed, is the exposure itself.

By exploiting the longitudinal data structure of the MCS, fixed effects models use within-unit changes in the outcome of interest over time (the fixed effects model effectively compares each child with herself/himself at an earlier time as her/his control) in order to account for difficult to measure individual-level factors that do not change with the passage of time and are likely to be associated with the exposure and the outcome of interest (e.g. gender, stable personality characteristics or innate ability). However, when changes in the exposure of interest are in the control of or they are associated with characteristics of the individuals under study (endogenous), there is a likelihood that unobserved time-varying factors that might be associated with changes in the exposure of interest (Oakes and Kaufman, 2017).

The difference-in-difference model complements the child fixed effects by using individual-level panel data and by leveraging factors external to individual that lead to changes in a child's exposure state. Individual changes in children's verbal skills and externalising or internalising symptoms in the presence of a change to attainment grouping are more likely to be plausibly unconfounded (the transition to attainment grouping is neither driven by outcomes nor unobserved confounders at the individual level) than changes in the absence of such school-level initiatives that could be attributed to a wide range of factors. In general, policy changes generate what are often referred to as ''natural experiments'' and form the typical setup for difference in difference analyses (Craig et al., 2012). By controlling for shared changes over time e.g. all schools in England and Wales went through the global financial crisis between 2008 and 2012, the difference-in-difference focused on the change to attainment grouping between MCS4 and MCS5 that occurred in some English and Welsh primary schools, but not others, and the relative changes in children's developmental outcomes. In simple words, a child fixed effects difference-in-difference modelling design can estimate a less biased effect of the transition to attainment grouping than conventional regression modelling by using changes in verbal skills for those children who changed to attainment grouping (treatment group) and those who did not (control group) (Meyer, 1995, Dimick and Ryan, 2014, Angrist and Pischke, 2009).

4.7.2: Growth Curve Modelling

To explore the long-term association between attainment grouping at age 7 and the development of MCS participants' externalising and internalising symptoms from age 7 years to 14 years (chapter 7), growth curve modelling was used. Growth curve modelling was not applied to explore the long-term association between attainment grouping at age 7 and trajectories of verbal skills from 7 to 14 years due to the use of different cognitive assessment instruments. At age 14 survey (MCS6), the standardised shortened version of the Applied Psychology Unit (APU) Vocabulary Test was used to measure MCS participants' verbal skills instead of the BAS which was used in previous MCS sweeps (Connely, 2013, Sullivan et al., 2017). Because both cognitive batteries are scaled in a different manner, this makes the combination and use of BAS and APU standardised scores difficult.

In recent years, growth curve modelling has gained popularity in various scientific fields; mainly among developmental psychologists (Flouri et al., 2015d, Flouri and Sarmadi, 2016, Flouri et al., 2016, Flouri and Midouhas, 2017, Flouri et al., 2017, Cillessen and Borch, 2006, Maulana et al., 2014, Sacker et al., 2005).

In growth curve modelling, all participants in each population are assumed to have developmental curves of the same functional form (e.g., linear, or curvilinear), but the parameters describing their curves may be different. For instance, with linear developmental curves, there may be individual differences in the initial level as well as in the growth rate or rate of change (Curran et al., 2010). Growth curve modelling is then a useful statistical technique to estimate these parameters. It is used to obtain a description of the mean growth in a population over a specific period. However, the main emphasis lies in explaining variability between subjects in the parameters that describe their growth curves, that is, in inter-individual differences in intra-individual change (Curran et al., 2010). The model on which growth curve modelling is based can be approached from several perspectives. For examining the relationship between early attainment grouping and the development of externalising or internalising symptoms from mid-childhood to mid-adolescence, the regression analysis model was constructed as a two-level multilevel regression (clustering at the ward level accounted for <5% of the variance in externalising or internalising difficulties , so it was ignored), where level 1 refers to observations to the lowest level of hierarchy (e.g. MCS participant's measurement on a given occasion i.e. occasion level), and level 2 refers to the cluster (MCS participant i.e. individual level).

A basic growth curve modelling is a combination of fixed effects and random effects that best capture the collection of individual trajectories over time. The biggest advantage of growth curve modelling or a mixed-effects model is that it can be used to investigate time invariant effects of the independent variables (here attainment grouping at age 7), whereas fixed effects models are designed to measure time variant characteristics (Gardiner et al., 2009).

In a growth curve modelling, a fixed effect represent a single parameter that exists in the population (e.g., the population mean externalising or internalising difficulties), and a random effect represents the random probability distribution around that fixed effect (e.g., the population variance in mean externalising or internalising difficulties). Consistent with these definitions, in the analyses of chapter 7 , the fixed effects represent the mean of the trajectory combining all the observations across MCS4 to MCS6 which correspond to 7, 11 and 14 years of age, and the random effects represent the differences of the individual trajectories around the average trajectory of parent-reported externalising or internalising difficulties. To put it clearly, for a linear trajectory, the fixed effects are estimates of the mean intercept (i.e., starting point) and mean slope (i.e., rate of change) that jointly define the underlying trajectory pooling of the entire sample; in contrast, the random effects are estimates of the between-individual variability in the intercepts and slopes. Smaller random effects (i.e., smaller variances of intercepts and slopes) imply that the parameters that define the trajectory are more similar across the sample of individuals. On the contrary, larger random effects (i.e., larger variances of intercepts and slopes) show greater individual differences in the magnitude of the trajectory parameters around the mean values; that is, some individuals are reporting higher or lower intercepts, or steeper or less-steep slopes relative to others. Taken together, the fixed and random effects capture the general characteristics of growth for the average child who participated in MCS 4, MCS 5 and MCS 6 and between-person differences around the average trajectories of externalising or internalising difficulties from age 7 to age 14 years. What follows is a description of the modelling approach followed for the scope of the statistical analyses of chapter 7.

4.7.2.1: Growth Curve Modelling strategy

Model 1: shows adjusted results by including main effects of attainment grouping at age 7, an interaction between attainment grouping at age 7 and time, MCS strata and gender. Information on MCS strata was included as a covariate in Model 1 in order to account for the stratified survey design of the MCS, since panel data analysis commands in STATA, such as *xtmixed* do not support the survey *svy* prefix in estimating the coefficients of interest. The rationale of this model was to assess the relationship between attainment grouping at age 7 and the rate of change of these initial differences between those who were grouped by attainment at age 7 and those who were taught in mixed attainment groups from 7 to 14 years, after accounting for gender and the disproportionately stratified design of the MCS.

Model 2: shows adjusted results for school-related confounders that were related to attainment grouping at age 7 and parent-reported symptoms of externalising or internalising difficulties at the same age. This was done to get closer to isolating the relationship between attainment grouping at age 7 and differences in the level of parent-reported externalising or internalising symptoms at the intercept (age 7), after adjusting for gender, the disproportionately stratified design of the MCS and school-related characteristics.

Model 3: a three-way interaction between attainment grouping at age 7, time and family income was added to model 2. This was done to explore whether attainment grouping at age 7 contributes to an increase in socioeconomic inequalities in parent-reported externalising and internalising symptoms over time, after adjusting for gender, the disproportionate stratified design of the MCS and baseline school-related characteristics.

Model 4: a three -way interaction between attainment grouping at age 7 years, time, and verbal skills at age 5 was added to model 2. This was done in order to explore whether attainment grouping at age 7 contributed an increase in psychological symptoms differences over time among MCS participants who entered UK primary school with a higher and lower level of verbal skills, after adjusting for gender, the disproportionate stratified design of the MCS and baseline school-related characteristics.

Because interaction tests in models 3 and 4 could be underpowered, these interaction tests were all considered as exploratory (Zammit et al., 2010).

All models compared findings from multiple imputed data (assuming MAR) estimates and complete cases to assess bias and efficiency of the growth curve modelling estimates when missing information was considered and when it was ignored.

Table 4. 12: Model summary (Chapter 7)

Model 1	Centred age (MCS sweep) [a]+ MCS stratum + attainment grouping at age 7 +attainment grouping at age 7 * age+ gender
Model 2	Model 2+ school-related confounders
Model 3	Model 2 + Family income+ Family income * attainment grouping at age 7 *age
Model 4	Model 2+ Verbal skills at age 5 +Verbal skills at age 5 * attainment grouping at age 7 *age
[a] Age was centred at MCS4 (age 7) [b] school type, class size at age 7, number of SEN pupils in child's class at age 7, number of EAL pupils in child's class at age 7, disruptive pupils in child's class at age 7	

4.8: Adjustment for sampling strategy

4.8.1: Clustering

The MCS dataset is hierarchically structured: children and parents are nested within families, then within wards. Therefore, the observations are not statistically independent. If this structure is ignored, standard errors of the coefficients could be underestimated, reducing precision and risking Type 1 error bias (Hansen et al., 2014).

Exploratory analyses revealed that ward-level variance was trivial for all models concerning verbal skills or socioemotional difficulties, so was ignored for simplicity. In the MCS family-level clustering only exists for multiple births. These children were excluded because of their small number and the possibility that their verbal cognitive and behavioural development is determined differently (Sutcliffe and Derom, 2006).

4.8.2: Stratification

As discussed earlier in this chapter, the MCS uses a disproportionate stratified sample (Plewis, 2007). Therefore, weighting of means, variances etc. should be considered to reduce sampling error. All descriptive analyses were conducted using STATA's survey command *svy*, thus accounting for MCS survey design (baseline MCS survey and attrition weights, stratification, and clustering). Statistically significant differences were assessed at the conventional 5% level ($p < 0.05$) (Spiegelhalter, 2019).

The model-based approach to statistical inference was adopted in the multivariable analyses of this PhD to adjust for disproportionate stratification in which families living in disadvantaged wards, English wards with high ethnic minority populations and wards in Wales, Scotland and Northern Ireland were deliberately over-represented. Stratification characteristics were included as covariates in models according to the recommendation of the CLS (Hansen et al., 2014) and because once the survey design is adequately represented in models it becomes ignorable, making sampling weights unnecessary (Skinner and Wakefield, 2017).

4.9: Treatment of Missing Data

In longitudinal studies it is also important to take into consideration non-response (Carpenter and Kenward, 2013). There are two types of non-response: unit non-response and item non-response. Unit non-response is comprised of attrition and sweep-non-response. Attrition is the permanent loss of participants over sweeps (monotone loss to follow-up), while sweep non-response refers to temporary loss of a participant at least in one sweep (non-monotone loss to follow-up) but returning to the study at later sweeps. Indeed, sweep non-response can be also considered a special case of item non-response where all variables are missing for the same respondent in a longitudinal record (Mostafa and Wiggins, 2015). Item non-response can take two forms; a) the complete absence for an interview or an assessment representing the child , for instance non-response or non-participation to a cognitive assessment at MCS5, but response to the child self-completion survey at the same wave, and b) the absence of answers to specific questions in an interview, for example refusing to respond to questions related to corporal punishment. Unit and item non-response can also be differential. MCS participants with different characteristics may have a higher probability of non-response. For instance, in the MCS, lone parents, young mothers, the less advantaged and ethnic minorities are more likely to have a higher probability of non-response (Hansen et al., 2014).

4.9.1: Missing data mechanisms

The missing data mechanism can be defined as the probability that data are missing, given the values of the observed and missing data. There are three explanations or mechanisms explaining why data are missing from observational studies.

When the probability that data are missing does not depend on the values of observed or missing data, then it is assumed that data are Missing Completely at Random (MCAR). For instance, when some data are accidentally lost due to technical issues.

Data are assumed to be Missing at Random (MAR) when the probability that data are missing depends on the values of the observed data but does not depend on the values of the missing data. Income data are MAR if for instance parents' reports of their income depends on other socio-economic characteristics such wealth.

Finally, data are assumed to be Missing not at Random (MNAR) when the probability that data are missing depends on the values of the missing data. Data on parental psychological distress are MNAR if respondents with high levels of psychological distress are more likely to have answered questions related to psychological distress. MNAR data can be also called as non-ignorable or informative missing data.

4.9.2: Multiple imputation

Multiple imputation was used to deal with missing data under MAR assumption. Missing data were imputed by applying the Multiple Imputation Chained Equations (MICE) algorithm. MICE was implemented in STATA version 15.1 using *mi impute chained* (Royston and White, 2011). MICE was the preferred multiple imputation method because it can handle different types of variables; each variable was then imputed using its own regression model. Where possible, recommendations on how to deal with sample design and non-response in analysing the MCS data in combination with multiple imputation were followed (Seaman et al., 2012). However, it was not possible to weight any of the advanced analyses (difference-in-difference) of this PhD since attrition weights could not be included when estimating fixed effects difference-in-difference or multilevel growth curve modelling. Detailed analyses of non-response in the MCS highlighted that non response does not differ substantially when the disproportionate stratification was included in the sampling design, so weighting adjustments for attrition are unlikely to substantially influence analyses (Plewis, 2007). This suggests that not weighting the data is unlikely to alter the results. A detailed description of the imputation models used for the analyses of this PhD is given in each results chapter.

4.9.3: Number of imputations

The number of multiple imputations needed was based on the recommendation that the resulting standard errors from the multiply imputed datasets should not be appreciably larger than their hypothetical minimum values (Enders, 2010). The magnitude of the multiple imputation standard error (more precisely sampling variance or squared standard error) relative to its theoretical minimum was calculated according to the Relative Efficiency formula. Relative Efficiency can be calculated as

$$\boldsymbol{Relative\ Efficiency} = \left(\mathbf{1} + \frac{\boldsymbol{FMI}}{\boldsymbol{m}}\right)^{-\mathbf{1}}$$

Equation 4. 1

Where m denotes the number of imputed datasets and FMI is the fraction of missing information (Rubin, 1987). A detailed account of multiple imputation models is provided in each empirical chapter of this thesis. All multiple imputed models have been informed by the formula of *Equation 4.1.*

The analyses of chapter 7 utilised teacher responses at MCS4 to explore the long-term association between attainment grouping at age 7 and the development of externalising and internalising symptoms from 7 to 14 years.

Chapter 5: Transition to attainment grouping and changes in children's verbal cognitive skills from 7 to 11 years.

5.1: Chapter overview

The aim of this chapter is to understand more about the relationship between the transition to attainment grouping between 7 and 11 years and the change in verbal skills over the same period. It was also assessed whether this relationship varied by family income or initial verbal skills. This chapter addresses objectives 1.1 to 1.3. Objective 1.1 examines whether the transition to attainment grouping between 7 and 11 years associates with the change in children's verbal skills over the same period. Objectives 1.2 and 1.3 test whether child's family income and baseline differences in verbal skills (age 5 years) moderate the association between the transition to attainment grouping between 7 and 11 years and the change in children's verbal skills over the same period. It is hypothesised that the transition to attainment grouping will be linked to an increase in socioeconomic inequalities in children's verbal skills and an increase in verbal skills gap from the baseline for children who transition from mixed to attainment groups between age 7 and 11 compared to those who remain in mixed groups. Firstly, analyses using MCS data, methods and results are presented. Then results are discussed together with reference to initial hypotheses and the previous literature. Table A2.3 in the Appendix 2 provides a detailed account of the variables used in the analyses of this chapter.

5.2: Methods

5.2.1: Multiple Imputation

5.2.1.1: Multiple imputation sample size

There was a higher response rate in teacher survey at MCS 5 than MCS 4. Therefore, the analysis sample was restricted to the number of responses given in the teacher survey at MCS5. Information was imputed for 7,196 singleton MCS participants.

5.2.1.2: Item non-response

Of the 7,196 MCS participants who were considered for multiple imputation, the greatest amount of missing information was observed as expected for the variables of MCS4 as it can be observed in Table 5.1.

Table 5. 1: Percentage of missing values in BAS verbal skills at ages 7 and 11 (N=7,196)

	Complete	Missing (N)	Missing (%)
Relationship between streaming and setting at MCS 4	4,082	3,114	43.27%
Relationship between streaming and setting at MCS 5	6,937	259	3.60%
Verbal skills at MCS 3	6,689	507	7.00%
Verbal skills at MCS 4	6,453	743	10.30%
Verbal skills at MCS 5	7,105	91	1.30%
Family income (MCS 5)	7,196	0	0%

5.2.1.3: Missing data mechanism and data imputation method

Missing information regarding attainment grouping, and verbal skills was assumed to be MAR (see Appendix 2-Tables A2.1 and A2.2). All missing values in the variables of interest were imputed by using Stata's MICE algorithm.

5.2.1.4: Imputation model specification

The imputation model included information on the transition to streaming or setting (for literacy or/and maths) variable, family income and verbal scores at ages 5, 7 and 11 years. Two interactions were included in the imputation model: a) between transition to attainment grouping variable and family income, and b) between transition to attainment grouping variable and baseline verbal skills (age 5 years).

Language spoken at home at the time of MCS survey and information on MCS participant's non-verbal reasoning ability (pattern construction and picture similarities at age 5 years) were included as auxiliary variables in the imputation model to improve the quality of the imputations. These variables were found to effectively predict missing information on verbal cognitive skills from age 5 to age 11.

Baseline survey and attrition weights were also included in the imputation models to consider the unequal probabilities of selection in the MCS and the differential loss to follow-up. The imputation model included information on MCS stratum to consider the stratified sampling design of the MCS and make sure that imputed values were not unduly influenced by subjects from groups that were oversampled in the MCS. According to calculations stemming from equation 4.1 formula (see chapter 4); Fifty (50) multiple imputed datasets were generated, and the reproducibility seed was set at 2003.

5.2.1.5: Imputation diagnostics

Statistical checks on the 50 multiple imputed datasets for BAS verbal scores revealed that MI produced reasonable values of imputed data (all imputed values on BAS scores ranged from 20 to 80). Further, the average within-imputation and between-imputation variances for all the other variables included in the imputation models were small and the relative efficiency of the imputed values exceeded 99%. There was also an acceptable amount of Monte Carlo Error (the Monte Carlo error on the coefficients of interest was <=10% of the standard error). This suggested that a sufficient number of imputations was used. In the analysis stage estimates from each imputed dataset were combined under Rubin's rules (Rubin, 1987) by using Stata's command *mi estimate*.

5.2.2: Analysis sample

Out of 14,392 observations from 7,196 MCS participants, thirty-six (36) observations from 18 children who did not have valid information on verbal skills assessments at least on one occasion at ages 5, 7 and 11 years were excluded from further empirical analyses. MCS participants who transitioned from attainment grouping at MCS4 (age 7) to no grouping at MCS5 (age 11-N=876), and those children already in attainment groups at both MCS4 and MCS5 (N=2,125) were omitted from further statistical analyses. Statistical analyses were conducted in 4,177 MCS participants using multiple imputed data and in 2,125 MCS participants without missing information on attainment grouping verbal skills from MCS 3 (age 5) to MCS 5 (age 11), and family income.

5.2.3: Analysis model specification

Fixed effects difference-in-differences modelling was applied in complete cases analysis (n=2,125) and multiple imputed data (n=4,177) in order to explore the association between the transition to attainment grouping between 7 and 11 years and the change in children's verbal skills at the same time period. Because it was assumed that unobserved variance was associated with the transition variable, fixed effects specification was used. Formal empirical comparison of fixed effects and random effects

[9] modelling using a Hausman test[10] showed that the use of random effects was inappropriate in this instance ($p<0.01$).

For this chapter's analyses, the 'treatment group' consists of those MCS participants who were not taught in either streams or literacy sets at age 7, but they changed to streaming or literacy setting between age 7 and age 11. The fixed effects difference-in-difference method takes account of observed changes in children's verbal skills as they grow older. Changes in children's verbal skills linked to the transition to attainment grouping between 7 and 11 years were adjusted to consider the change in children's verbal skills who were not grouped by attainment when they were 7 and 11 years old. The fixed effects difference-in-differences also considered the temporal changes in children's verbal skills that would have been overestimated if simple differences were calculated. The effect of the transition to attainment grouping on changes in verbal skills from 7 to 11 years were calculated according to equation 5.1

$$\boldsymbol{Y_i g_t} = \boldsymbol{\gamma_0} + \boldsymbol{\gamma_1 T_g} + \boldsymbol{\gamma_2 P_t t} + \boldsymbol{\gamma_3} \left(\boldsymbol{T_g} \times \boldsymbol{P_t}\right) + \boldsymbol{f_{it}} + \boldsymbol{g_t}$$

Equation 5. 1

Where:

$\boldsymbol{Y_i g_t}$ =average verbal cognitive skills for the i^{th} child during study period t

$\boldsymbol{\gamma_0}$= Baseline average

---

[9] In random effects modelling, it is assumed that unobserved characteristics are not correlated with the exposure of interest. If the assumption does not hold, the association between the exposure and the outcome will be biased.

[10] The Hausman test can empirically evaluate the presence of correlation between unobserved characteristics and the exposure of interest by comparing fixed effects to random effects estimates through a global test.

$\boldsymbol{\gamma_1 T_g}$ = Time trend in the control group

$\boldsymbol{\gamma_2 P_t\, t}$ =Difference between control and treatment group before the intervention.

$\boldsymbol{\gamma_3}\ (\boldsymbol{T_g} \times \boldsymbol{P_t})$ =The difference in changes over time.

$\boldsymbol{f_{it}}$ is the child fixed effect for the i^{th} child capturing time invariant child characteristics during study period t.

$\boldsymbol{g_t}$ = Time-varying error

All the difference-in-differences models were estimated with robust standard errors to account for heteroscedasticity and autocorrelation between pre and post change in verbal skills in the same individual (Wooldridge, 2013).

The possibility of heterogeneous effects was explored by extending the model of *equation 5.1* to include a third variable (either family income or baseline verbal), forming accordingly a difference-in-difference-in-difference or a triple differences approach. When considering triple difference models, all main effects and the three two-way interactions (lower-level interaction terms) must also be included so that the model remains hierarchically well specified (Oakes and Kaufman, 2017). Caution should be exercised when interpreting these three-way interactions since these interactions could be underpowered (Heo and Leon, 2010), thus these interaction tests should be perceived as exploratory.

Statistical inferences in difference-in-differences settings involve multiple comparisons by including multiple interaction terms; however, multiple comparison adjustments/corrections to avoid false discoveries are not needed when the structure of the data is hierarchical (multilevel) (Gelman et al., 2012). Particularly, multilevel models compared to classical regression modelling perform partial pooling, that is shifting point estimates toward each other and their corresponding confidence intervals closer to each other, whereas classical methods alter p-values or make confidence intervals wider (Gelman et al., 2012). Therefore, the hierarchical structure of the data used to perform difference-in-difference regression modelling in this chapter could effectively address the issue of multiple comparisons and subsequent false discoveries and produce more efficient estimates.

5.2.3: Model assumptions

The fixed effects difference-in-differences approach makes the identifying assumption that the control group serves as an adequate proxy for the counterfactual outcome that would have been observed in the treatment group had they not been treated (Ryan et al., 2015). With only two time points available, the identifying assumption of the difference-in-differences rests on parallel growths which implies that in the absence of the transition to attainment grouping, the treated and the controls would have had parallel growth paths equal to the output levels in first differences (before treatment period) (Reggio and Mora, 2012).

An additional, but equally important identifying assumption of the difference-in-differences modelling is the absence of unobserved time-varying confounding associated only with the group of children who transitioned to attainment grouping and changes in their verbal skills.

5.3: Results

5.3.1: Distribution of child verbal scores at MCS 4 and MCS 5.

The children's verbal scores at ages 7 and 11 years were normally distributed as shown in Figure 5.1. There was also a shift in the distribution of verbal cognitive scores at age 11 to the right of the distribution of scores at age 7, highlighting an increase of verbal scores over time.

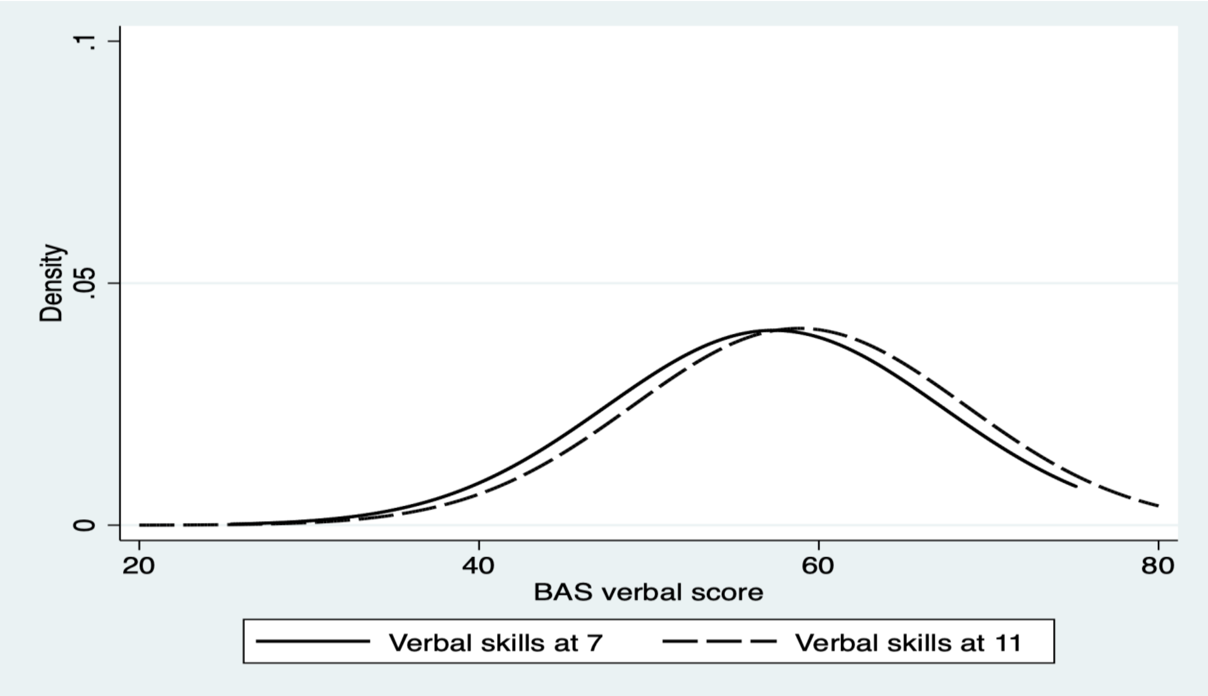

Figure 5. 1: *Distribution of BAS verbal cognitive scores at ages 7 and 11years (N=4,177)*

Figure 5.2 shows that there was a small, but statistically significant increase in children's mean verbal cognitive scores from age 7 to age 11 years.

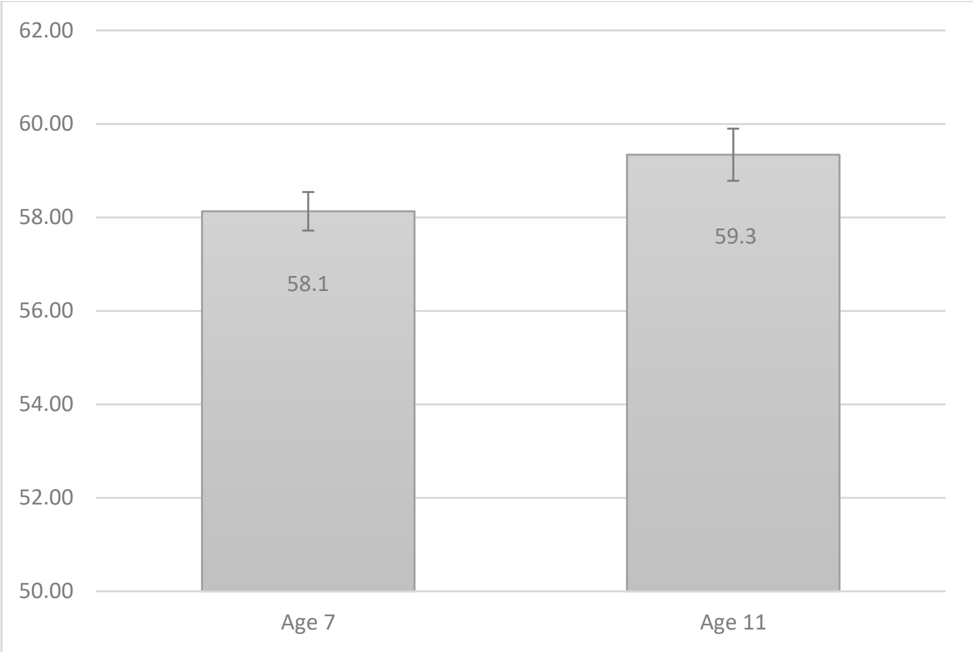

Figure 5. 2: Average BAS verbal cognitive scores at ages 7 and 11 years (N=4,177)

5.3.2: Fixed effects difference-in-differences results

5.3.2.1: Transition to attainment grouping and children's verbal skills between ages 7 and 11 years

Table 5.2 and Figure 5.4 illustrate a significant increase in verbal scores from age 7 to age 11 by approximately 1.7 BAS points (from a BAS score of approximately 57.5 to a BAS score of 59) among MCS participants who did not change to attainment grouping between ages 7 and 11 years, using both the complete cases analysis and multiple imputed data. MCS participants who changed to attainment grouping between ages 7 and 11 years experienced similar gains in their verbal scores across the same period to those MCS participants who did not change.

Table 5. 2 The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores

	Complete case analysis (n=2,125)				Multiple imputation (n=4,177)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.37***	0.23	0.92	1.81	1.65***	0.237	1.19	2.12
Attainment grouping * MCS5	-0.2	0.37	-0.93	0.53	-0.19	0.35	-0.88	0.50

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** p<0.001, ** <0.01, * <0.05

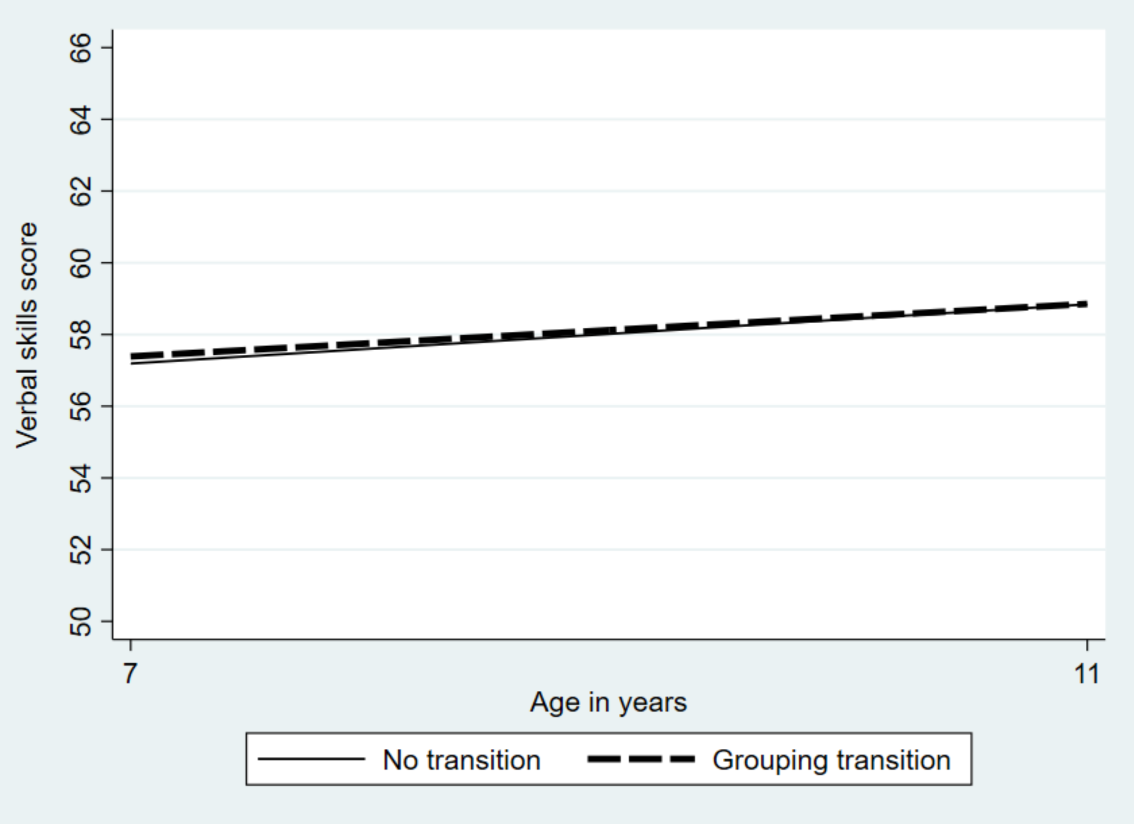

Figure 5. 3: Transition to attainment grouping and change in verbal skills

5.3.2.2: Transition to attainment grouping and socio-economic inequalities in verbal development?

Table 5.3 and Figure 5.5 show the results of a three-way interaction between family income, transition to attainment grouping and MCS 5. MCS participants in the highest family income group who transitioned from mixed to attainment grouping experienced a similar improvement in their verbal skills to MCS participants who did not change to attainment grouping between ages 7 and 11 years. MCS participants at the lowest family income group who transitioned to attainment grouping had a slightly higher verbal score at age 7 (a BAS verbal score slightly above 54) compared to MCS participants at the lowest family income group who did not change to attainment grouping (a BAS verbal score slightly above 53). MCS participants in the lowest family income group who transitioned to attainment grouping ended up having a similar verbal score at age 11 (a BAS verbal score a little bit above 54) compared to MCS participants who did not change to attainment grouping . Although the small increase in socioeconomic inequalities between MCS participants in the highest and the lowest family income group who transitioned to attainment group, this increase was not statistically significant at the 5% level (coefficient -1.42, 95% CI -3.62, 0.78).

Table 5. 3: The association between transition to attainment grouping between 7 and 11 and changes in children's verbal cognitive scores between 7 and 11 by family income

	Complete case analysis (n=2,125)				Multiple imputation (n=4,177)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS4 to MCS 5	1.37	0.17	-0.57	3.32	0.76	0.45	-0.11	1.64
Attainment grouping * MCS5	-0.36	0.76	-2.67	1.94	0.25	0.63	-0.98	1.48
Highest incomes * MCS5 (Ref.)								
2nd quantile * MCS5	-1.11	0.48	-4.18	1.96	0.24	0.77	-1.26	1.75
3rd quantile * MCS5	1.61	0.27	-1.24	4.46	0.87	0.77	-0.64	2.38
4th quantile * MCS5	2.70	0.08	-0.29	5.69	2.19**	0.83	0.57	3.81
Lowest incomes * MCS5	-0.80	0.61	-3.82	2.23	0.88	0.84	-0.76	2.52
Attainment grouping * Highest incomes * MCS5 (Ref.)								

	Complete case analysis (n=2,125)				Multiple imputation (n=4,177)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Attainment grouping * 2nd quantile * MCS5	0.87	0.64	-2.82	4.57	0.48	1.07	-1.62	2.58
Attainment grouping * 3rdquantile MCS5	-0.46	0.80	-3.94	3.03	0.67	1.10	-1.48	2.82
Attainment grouping* 4th quantile * MCS5	-2.79	0.15	-6.59	1.01	-1.16	1.13	-3.39	1.06
Attainment grouping * Lowest incomes * MCS5	0.76	0.71	-3.29	4.81	-1.42	1.12	-3.62	0.78

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

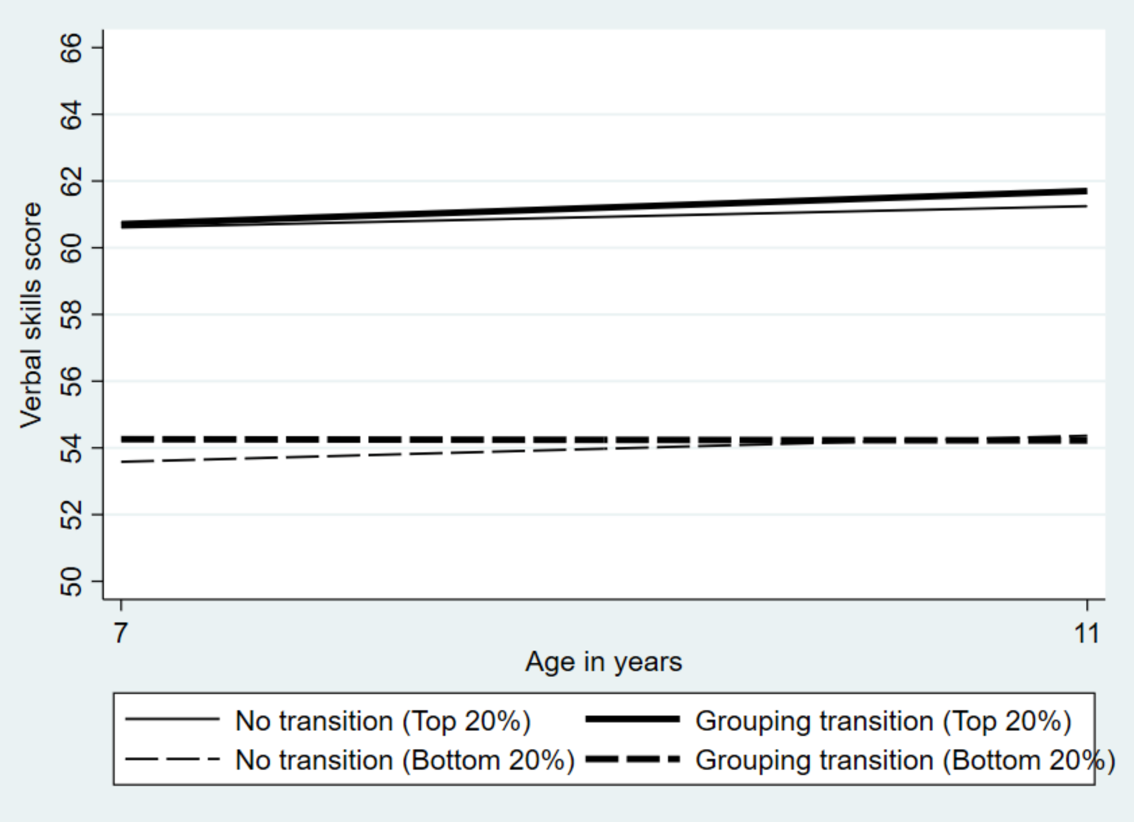

Figure 5. 4 :Transition to attainment grouping and change in verbal skills by MCS participant's family income

5.3.2.3: Transition to attainment grouping and verbal development by baseline verbal skills

Table 5.4 and Figure 5.6 shows a three-way interaction term between transition to attainment grouping, time, and verbal skills at age 5. MCS participants with higher verbal skills (here the top 20%) at age 5 who transitioned from mixed to attainment grouping experienced a similar improvement in their verbal skills to MCS participants who did not change to attainment grouping between ages 7 and 11 years. MCS participants with lower verbal skills (here the bottom 20%) at age 5 who transitioned to attainment grouping had a higher verbal score at age 7 (a BAS verbal score close to 53) compared to MCS participants at the bottom 20% of the verbal skills distribution at age 5 who did not change to attainment grouping (a BAS verbal score of 52). Nevertheless, MCS participants at the bottom 20% of the verbal skills distribution at age 5 age 5 who transitioned to attainment grouping improved at a slower rate between MCS 4 and MCS

5 compared to those ones who did not change to attainment grouping (a change in BAS verbal score from 52 to 54). Table 5.4 highlights that the three-way interaction was not statistically significant at the 5% level However, it is plausible to argue for an indication of an increase in verbal skills differences from the baseline (MCS 3-age 5) between children who changed to attainment grouping between MCS 4 and MCS 5 and those who did not.

Table 5. 4: The association between transition to attainment grouping between 7 and 11 and changes in children's verbal cognitive scores between 7 and 11 by baseline verbal cognitive skills (age 5)

	Complete case analysis (n=2,125)				Multiple imputation (n=4,177)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	2.47	1.259	0.00	4.94	2.44*	1.227	0.03	4.85
Attainment grouping * MCS5	-3.41	2.053	-7.43	0.62	-2.92	1.74	-6.35	0.51
Baseline verbal cognitive skills* MCS5	-0.02	0.022	-0.06	0.02	-0.01	0.021	-0.06	0.03
Baseline verbal cognitive skills* Attainment grouping * MCS5	0.06	0.036	-0.01	0.13	0.05	0.03	-0.01	0.11

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

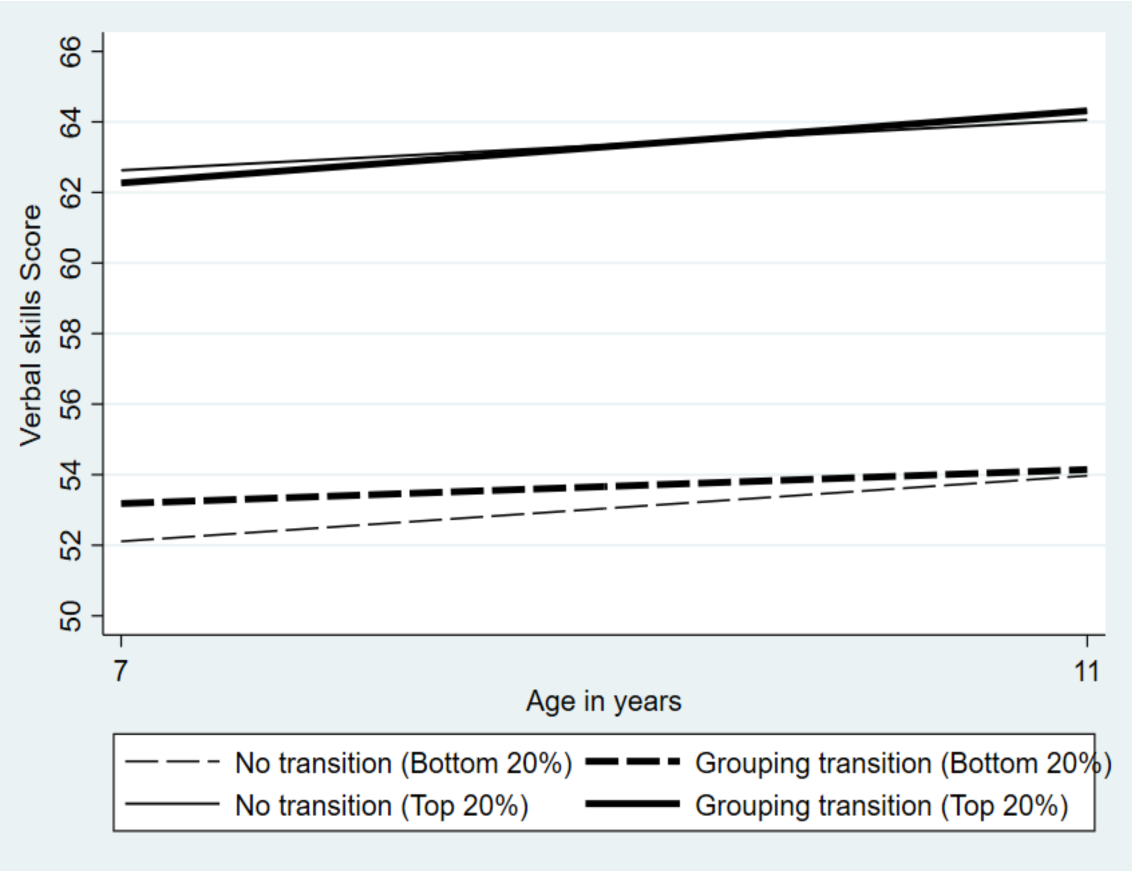

Figure 5. 5: Transition to attainment grouping and change in verbal skills by MCS participant's baseline verbal skills

5.3.3: Sensitivity analyses

To explore whether the findings of this chapter were robust under additional assumptions, sensitivity analyses were conducted using complete cases only since no considerable differences in the coefficients of models using complete cases and imputed data were observed.

5.3.3.1: Adjusting for school-related characteristics related to a transition to attainment grouping

Available MCS time-invariant or school-related characteristics that do not change much over short time and were associated with a transition to attainment grouping between MCS 4 and MCS 5 (see Table 4.10) were included as controls to account for school-level differences in children’s average verbal skills and their association with differential verbal skills change/progress during KS2.

Tables A2.4 to A2.6 in Appendix 2 show that adjustment for the number of classes in the child’s school year, class size, number of children with EAL in the child’s classroom and whether the child was taught in a single or mixed year classroom did not change the results of this chapter.

5.3.3.2: Adjustment for schools moves between MCS 4 and MCS 5

Attainment grouping experiences may differ between MCS 4 and MCS 5 among children who changed primary school and those who did not during the same period. To consider the possibility of differential changes in developmental outcomes among children who changed to attainment grouping, a variable indicating primary school moves between MCS 4 and MCS 5 was added as an additional control.

The results of tables A2.7 to A2.9 in Appendix 2 highlight that statistical control for school moves between MCS 4 and MCS 5 did not alter the results of this chapter.

5.3.3.3: Transition to streaming/literacy setting between MCS 4 and MCS 5 and change in children’s verbal skills

The pedagogical practices and the content of lessons may substantially vary between setting for maths and streaming and/or literacy setting. Setting for maths is less likely to be associated with children’s verbal skills than streaming and literacy setting. To consider a possible dilution of the regression coefficients due to inclusion of setting for maths, fixed effects difference-in-difference modelling re-examined the association between transition to streaming/setting for literacy between MCS 4 and MCS 5 and the change in children’s verbal skills over the same period.

Table A2.12 demonstrates that the increase in baseline verbal skills differences was more evident when the transition to streaming/literary setting was only considered (coefficient 0.10, 95% CI 0.02, 0.19, p=0.015-Table A2.12 vs. coefficient 0.06, 95% CI -0.01, 0.13, p=0.109 when setting for maths was included -Table 5.4). Therefore, taking into consideration the transition to maths setting between MCS 4 and MCS 5 is likely to reduce the influence of the transition to streaming or/and setting on literacy.

5.3.3.4: The influence of attainment grouping placement on inequalities in children's verbal skills between MCS 4 and MCS 5

Inequalities in children's verbal skills associated with a transition to attainment grouping may be influenced by children's placement to the top, middle or bottom attainment groups. To explore this possibility, the transition to top, middle, and bottom streams and sets was considered.

Table A2.14 show that adjustment for placement at the top, middle or bottom streams or sets for literacy and maths between MCS 4 and MCS 5 resulted in a stronger association between transition to attainment grouping (streaming or/and setting for literacy or/and setting for maths) and an increase in children's verbal skills differences from MCS 3 (coefficient 0.13, 95% CI 0.03, 0.22, p=0.008), but not in the change of socio-economic inequalities in children's verbal skills (see Table A2.13). This finding highlights that social comparisons may take place between higher and lower attaining students not only between top and lower groups, but also within top and lower groups. As a result, a divergence of verbal cognitive attainment during KS2 is likely. In general, the verbal skills gap between students at the top and the bottom distribution of verbal scores at MCS 3 grew by 4.35 BAS points (95% CI 1.17, 7.54, p=0.007) or by as little as 10% to 70% of a standard deviation, with a mean estimate of approximately 40% of a standard deviation. Thus, it can be argued that there is a strong association between transition to attainment grouping (streaming/setting for literacy/setting for maths) and an increase in children's verbal skills differences, independent of their allocation to top, middle and bottom streams as well as sets for literacy and maths.

5.3.4: Summary of findings

***Objective* 1.1: To examine whether a transition from mixed attainment grouping to an attainment grouping (streaming/setting for literacy) between ages 7 and 11 years is related to progress in children's verbal skills.**

MCS participants who transitioned from mixed attainment grouping to attainment grouping (streaming/setting for literacy/setting for maths) between MCS 4 and MCS 5 years did not exhibit significantly slower improvement in verbal skills compared to children who were taught in mixed attainment groups at both ages 7 and 11.

Objective 1.2: To test whether a transition from mixed to attainment grouping between ages 7 and 11 years is related to a change in socioeconomic inequalities in children's verbal skills between ages 7 and 11 years.

Limited evidence in support of the hypothesis 1.2 was found. The transition from mixed to attainment grouping (streaming/setting for literacy or/and maths) between ages 7 and 11 years was not significantly associated with an increase in socio-economic inequalities in verbal skills from age 7 to age 11. Socio-economic inequalities in children's verbal skills were similar at both time points, regardless of a transition to streaming/setting for literacy/or and maths.

***Objective 1.3:* To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a change in the gap in verbal skills between age 7 and 11 contingent on the gap in verbal skills at primary school entry (age 5).**

The gap in verbal skills at school entry (high versus low scores) did not significantly widen for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups. However, controlling for a transition to top, middle, and bottom streams, literacy and maths sets resulted in a significant increase in verbal skills differences from the baseline among those who changed to attainment grouping. In addition, a significant increase in verbal skills differences from the baseline was observed for children who changed to streaming/literacy setting only.

5.4: Discussion

It has been demonstrated in the literature that there is little or no effect of attainment grouping on raising academic standards in UK secondary schools. I hypothesised that children who transitioned to an attainment grouping between age 7 (MCS 4) and age 11 (MCS 5) would experience a smaller improvement in their verbal skills than children who did not change to attainment grouping. Findings from the fixed-effects difference-in-differences did not support this hypothesis. Also, little evidence was found to support a differential effect of the transition to attainment grouping on children's verbal skills growth according to family income. What is striking is that children from lower-income families who either changed to attainment grouping or were taught in mixed-attainment classes at both MCS 4 and MCS 5 experienced no or very little improvement in their verbal skills. The hypothesis that an increase in verbal cognitive inequalities in response to the transition to attainment grouping between age 7 and 11 years was supported (hypothesis 1.3). Transition to attainment grouping was associated with an increase in children's verbal skills differences from the baseline. Findings are discussed below according to the existing literature.

5.4.1: Comparison to previous literature

With respect to hypothesis 1.1, no differential (slower) improvement in children's verbal skills was observed for those children who changed to attainment grouping between 7 and 11 years and those who did not in England and Wales. However, this finding is difficult to compare with previous published studies as the focus of these studies has often been on the role of attainment grouping on performance (Key Stage tests or high stakes tests) and not on aptitude test scores (low-stakes tests), primarily among secondary aged pupils.

This is the first attempt to investigate the effect of the change to attainment grouping on subsequent changes in verbal skills in primary school aged children in a contemporary UK context. In this respect, this finding is novel and useful in enriching the debate around the suspected benefits or harms of attainment grouping in primary school pupils' verbal cognitive development. This finding is consistent with early findings from Kutnick and colleagues (2006). The authors of this study concluded that primary schools which adopted setting rarely had higher KS 2 levels in English than their local authority and national averages. This finding also corroborates recent findings from a second-order meta-analytic review which examined the effect of between-class attainment grouping on achievement test scores from Kindergarten to Year 12 in the US. Specifically, Steenbergen-Hu and colleagues (2016) reviewed evidence over the past 100 years on the role of grouping by attainment (between-class grouping) and found a non-significant effect of between-class attainment grouping practices on achievement test scores among primary and secondary education pupils in the US.

This chapter's analysis provided little evidence in support of differential verbal skills improvement between the least and the most advantaged MCS participants who were grouped by attainment between MCS 4 and MCS 5.

A key finding to emerge from the analyses of this chapter, is the differential effect of the transition to attainment grouping on the development of verbal skills among children who entered primary school in England and Wales with different levels of verbal skills. This is the first empirical attempt in a contemporary UK context to assess whether attainment grouping in primary school exacerbates initial verbal cognitive differences. This finding supports the divergence hypothesis that was outlined in chapter 2 (see section 2.7), of increasing differences in attainment during primary schooling among those who were taught in attainment groups. It can also be argued that it is consistent with recent Social Mobility Commission reports which highlighted that attainment grouping mainly in the secondary school setting hindered the progress of students who attained low scores in their KS2 exams (Social Mobility Commission, 2017), thus increasing inequalities in academic attainment. However, this finding contradicts recently published results of a second-order meta-analysis from (2016) which found no differential effect of between-class attainment grouping in the US primary and secondary education on students' academic attainment by initial attainment levels or learning abilities. A possible explanation for this contradictory finding might be the timing of publication of the included reviews; the earliest primary study in this second-order meta-analytic review by (2016) was conducted in 1922 and the latest in 1994, with the majority of the studies appearing between 1960 and 1970. Therefore, it is plausible to assume that the nature of education in the US, as well as the UK, has changed radically, since many of the studies were conducted in terms of generation or cohort effects, and their findings cannot be extrapolated to the academic/cognitive outcomes of the recent generation or cohort of students either in the US or the UK.

Still, given this contradiction, it is important to consider what might be explaining the increase in verbal inequalities during primary school in a contemporary English and Welsh context as a response to the transition to attainment grouping than taking this increase at face value. There are a couple of possible explanations on why the transition to attainment grouping might be candidate trigger for the observed increase in children's verbal cognitive differences in the MCS.

Firstly, placement of students with lower attainment at primary school entry into attainment groups later in primary school, especially placement to lower (middle and bottom) attainment groups, may result in low academic and general self-concept. In turn, this can lead low attaining students in these groups reducing their emphasis on academic aspects of their schooling and placing a greater value on social and physical aspects (McManus, 2010). The increase in children's verbal skills gap from the baseline (MCS 3-age 5) among those who were grouped by attainment between MCS 4 and MCS 5 (KS2) may also be the result of social comparisons. Low attaining students in lower streams or sets may have reduced academic self-concept through "Big fish-little pond" effect processes (Marsh, 1984), that is by making less favourable comparisons of their academic potential compared to students with comparable academic attainment in mixed-attainment groups.

Secondly, the increase in student's initial verbal skills differences related to the transition to attainment grouping may be due to differential teacher experience, teacher expectations, and pedagogical practices within attainment groups. There is some evidence that teachers who have lower qualifications are less experienced tend to be allocated to bottom streams or sets by their school's senior leadership team. Teachers who are perceived as being of "higher quality" or more experienced tend to be allocated to top streams or sets (Slavin, 1990, Ireson and Hallam, 2001). Such decisions have significant consequences especially for low-attaining students who are most likely to end up in middle and bottom streams or sets. There is also evidence that students who are placed to bottom streams or sets have a higher probability of experiencing more changes of a teacher (Kutnick et al., 2005, Ireson and Hallam, 2001, Francis et al., 2019). Teacher expectations and related pedagogy might also be different for higher or lower attaining students who are allocated to top streams or sets than those who are allocated to bottom streams or sets. There is evidence that teachers respond differently according to the level of 'ability' of their students (Francis et al., 2019). An association exists between placement to top streams or sets and higher teacher expectations through the provision of more challenging and fast-paced school-work, while being a low-attaining student in a bottom stream or set is associated with less challenging and slow-paced school work that covers a smaller part of the curriculum (Boaler et al., 2000). Academic research shows that students in top streams or sets are more likely to be given more homework by their teachers than students in bottom streams or sets (Ireson and Hallam, 2001). In addition, teachers might adopt a smaller variety of teaching styles when they teach students within attainment groups than children who are not taught in attainment groups, independent of academic/cognitive attainment level (Boaler et al., 2000, Ireson et al., 2005). It is therefore likely that lower attaining students at primary school entry who are later allocated to top streams or sets cannot cope with the challenging and fast-paced schoolwork of their stream or set. This may fulfil them with doubts about their academic capacity compared to their higher attaining peers in the allocated stream or set, which in turn may negatively influence their academic performance.

Finally, the very small verbal skills improvement of MCS participants who started primary school in England and Wales with a lower level of verbal cognitive skills and who were grouped by attainment between MCS 4 and MCS 5 might also be due to the differential formation of student identities (Kutnick et al., 2005). The transition to attainment grouping for lower attaining students at primary school entry, and their likely placement to middle and bottom streams or literacy sets, might have damaged their self-perception of what they were able to do, and it might have reduced their confidence. It might have also impacted on their school life and the adoption of alternative forms status (anti-social behaviour at school) with detrimental effects on their cognitive attainment and aspirations (Ball, 1981).

5.4.2: Strengths and limitations

In this chapter, complete cases results were compared to findings using multiple imputed data. For analysis with a larger sample size, the magnitude of estimates was reduced slightly, indicating that coefficients may be overestimated in smaller samples. Estimates for the complete cases showed the same level of significance with multiple imputation, although the standard errors were higher in complete cases than multiple imputed data. Part of the slight differences observed for the complete cases and multiple imputed results could be due to differences in sample characteristics. From the substantive analysis in this chapter, the complete cases sample was composed of MCS participants from more advantaged socially and economically as highlighted in the analyses of chapter 4.

This chapter's analyses have also established a quantitative framework for examining the causal effect of the transition to attainment grouping during primary school in the UK, as well as the existence of differential effects on children's development of verbal skills. The application of fixed effects difference-in-differences approach provided limited room for omitted variable bias (confounding). Causal inference based on observational data lies on finding an adequate substitute population for the unobservable counterfactual of interest in the absence of random allocation (Hernan and Robins, 2006, Maldonado and Greenland, 2002). This study devised a control group (those children who were not grouped by attainment at age 7 and age 11) to explore what would have been observed in the development of verbal cognitive skills of those children who changed to attainment grouping if, counter to the fact, they had not changed to attainment grouping.

This study also has important limitations which should be taken into consideration. The applied fixed effects difference-in-differences design has utilised observational data. There is always the possibility of unmeasured differences between the group of children who changed to attainment grouping (treatment group) and the group of children who did not change to attainment grouping (control group) that are not controlled with the difference-in-difference design (e.g the existence of confounding factors that vary over time for some schools in the treatment group, but not in the control group).

The results of the fixed effects difference-in-difference-in-differences should also be interpreted with caution. Interaction terms always have reduced power in statistical tests compared to main effect terms, meaning that the significance of the three-way interactions should be considered with caution as they might be prone to high rates of type I[11] and type II[12] error (Greenland, 1983, Heo and Leon, 2010).

---

[11] Type I error is the rejection of a true null hypothesis (also known as a "false positive" finding or conclusion).

[12] Type II error is the non-rejection of a false null hypothesis (also known as a "false negative" finding or conclusion).

In summary, this chapter provided little evidence of an association between the change to attainment grouping between ages 7 and 11 years and a smaller improvement the change in MCS participants' verbal skills over the same period. There was also little evidence that socio-economic differences in verbal skills increased in primary schools in response to attainment grouping transition, while there was some evidence of an increase in children's verbal skills differences during primary school among those who changed to attainment grouping. The next chapter will focus on the association between the transition to attainment grouping and changes in MCS participants' externalising or internalising psychological symptoms from age 7 to age 11.

Chapter 6: Attainment grouping transition, externalising, and internalising symptoms from 7 to 11 years.

6.1: Chapter overview

This chapter contains statistical analyses related to objectives 2.1 to 2.3. Objective 2.1 will examine whether the transition to attainment grouping between 7 and 11 years is associated with a change in children's externalising or internalising symptoms during the same period. It is hypothesised that being exposed to a transition from mixed grouping to attainment grouping will be associated with a smaller decrease in parent- or teacher-reported externalising psychological symptoms and with a bigger increase in parent- or teacher-reported internalising psychological symptoms during the same period.

Objective 2.2 explores whether the association between the transition to attainment grouping and the change in children's externalising or internalising symptoms years varies by family income. It is hypothesised that a transition from mixed to attainment grouping between will be associated with an increase in socio-economic inequalities in externalising and internalising psychological symptoms from age 7 to age 11. The gap in externalising and internalising psychological symptoms between the most advantaged and the least advantaged MCS participants will be wider by the end of primary school among MCS participants who transitioned from mixed to attainment grouping. The gap in externalising and internalising psychological symptoms between the most and the least advantaged MCS participants who are taught in mixed attainment groups at both ages 7 and 11 years will increase by less.

Objective 2.3 explores whether a transition from mixed grouping to attainment grouping is associated with an increased gap in externalising and internalising psychological symptoms from age 7 to age 11 between MCS participants with higher verbal skills and participants with lower verbal skills at age 5 (MCS 3). It is hypothesised that the gap in externalising and internalising psychological symptoms for MCS participants with different levels of verbal skills at school entry (high versus low scores) will widen among those who transition from mixed to attainment groups between MCS 4 and MCS 5compared to those who remain in mixed groups.

Firstly, methods specific to this chapter are described and results are presented. Results are discussed together with reference to initial hypotheses and previous literature. A detailed description of the variables used for the empirical analyses of this chapter is given in Table A3.5 in the Appendix 3.

6.2: Multiple Imputation

6.2.1: Imputation model sample size

The response rate to the teacher survey in England and Wales was higher in MCS 5, leading to a larger usable sample compared to the MCS 4, even after attrition. Therefore, multiple imputation was conducted using information from 7,196 singleton MCS participants at MCS 5.

6.2.2: Item non-response

Of the 7,196 MCS participants who were considered for multiple imputation, Table 6.1 shows that the greatest amount of missing information was observed for teacher reported SDQ (38.9%) and the attainment grouping variable at age 7 (43.3%).

Table 6. 1: Percentage of missing values in parent and teacher reported SDQ at 7 and 11, transition to attainment grouping and family income (N=7,196)

	Complete	Missing (N)	Missing (%)
Attainment grouping at MCS 4	4,082	3,114	43.3%
Attainment grouping at MCS 5	6,937	259	3.6%
Teacher-reported externalising score at MCS 4	4,396	2,800	38.9%
Teacher-reported externalising score at MCS 5	7,032	164	2.3%
Teacher-reported internalising score at MCS 4	4,395	2,801	38.9%
Teacher-reported internalising score at MCS 5	6,991	205	2.8%
Parent-reported externalising score at MCS 4	6,477	719	10.0%
Parent-reported externalising score at MCS 5	6,982	214	3.0%
Parent-reported internalising score at MCS 4	6,472	724	10.1%
Parent-reported internalising score at MCS 5	6,988	208	2.9%
Family income	7,196	0	0%
Verbal skills at MCS 3 (age 5)	6,689	507	7.0%

6.2.3: Missing data mechanism and imputation method

The MICE algorithm was implemented in Stata version 15.1 using *mi impute chained* since missing data in parent-reported and teacher-reported SDQ were assumed to be MAR (see Tables A3.1 to A3.4 in the Appendix 3).

6.2.4: Imputation models specification

Two separate imputation models were formed for parent-reported and teacher-reported SDQ. The two imputation models included information on the attainment grouping transition variable, on either parent-reported or teacher-reported externalising and internalising symptoms at MCS 4 and MCS 5 and school type. Interactions were also included in the two distinct imputation models: a) between the attainment grouping transition variable and family income and b) between attainment grouping transition variable and verbal skills at MCS 3.

Language spoken at home at the time of the survey and pattern construction at MCS 3 were used as auxiliary variables in the imputation models in order improve the accuracy of prediction of missing information on parent-reported and teacher-reported SDQ at MCS 4 and MCS 5. Preliminary analyses found that these three auxiliary variables were strong predictors of response to parent SDQ and predicted response to teacher SDQ.

Information on MCS stratum and survey weights, as well as an interaction between the two variables were also included in the imputation models. Fifty (50) multiple imputed datasets were generated for both parent-reported and teacher-reported child outcomes and the reproducibility seed was set at 2301 and 2302 respectively.

6.2.5: Multiple imputation diagnostics

Statistical checks using Stata's *midiagplots* command (Eddings and Marchenko, 2012) on the 50 multiple imputed datasets for parent-reported SDQs and the 50 multiple imputed datasets for teacher-reported SDQs, revealed that multiple imputation produced reasonable values of imputed data (all imputed values for parent and teacher reported externalising or internalising scores lied between 0 and 20).

The average within-imputation and between-imputation variances for all the other variables included in the imputation models were small and the relative efficiency of the imputed values exceeded 99%. There was also an acceptable amount of MCE. In the analysis stage estimates from each imputed dataset were combined under Rubin's rules (Rubin, 1987) by using Stata's command *mi estimate*.

6.3: Analysis sample

Fixed-effects difference-in-difference modelling was applied in 2,031 and 2,190 MCS participants with complete information on attainment grouping transition variable, parent- and teacher-reported SDQ respectively (complete cases). Fixed effects difference-in-differences modelling was applied using multiple imputed data on 4,157 and 4,181 MCS participants respectively with observations the grouping transition variable and parent- or teacher-reported SDQ, and all other analyses' variables. The analytic samples were derived after excluding observations of MCS participants who were taught in attainment groups at both age 7 (MCS 4) and (MCS 5), those who transitioned from attainment grouping at MCS 4to mixed teaching at MCS 5 , and those with missing information on parent-reported (n=51) and teacher-reported SDQ (n=27) at age 7 and age 11. Provided the time and funding constraints for the completion of this PhD, information of MCS participants who changed from attainment grouping at MCS 4to no grouping at MCS 5 and exclusion of those were grouped by attainment at both MCS 4 and MCS 5 was not considered for the fixed-effects difference-in-differences modelling.

6.4: Model specification

Provided that there was available information on attainment grouping for just two observation time points in the MCS, a basic fixed effects difference-in-differences design was adopted where outcomes are observed for two groups for two time periods. In the analyses of this chapter, the treatment group consists of those MCS participants who were not taught in attainment groups at age 7 but were subsequently changed to be taught in attainment groups at MCS 5. This method takes account of observed changes in children's parent- or teacher-reported symptoms of externalising and internalising difficulties as they get older. Changes in children's symptoms of externalising or internalising difficulties linked to the transition to attainment grouping are adjusted to consider the change in children's externalising or internalising symptoms who were not grouped by attainment at both age 7 and age 11 (control group). Therefore, this method considers the temporal changes in children's externalising and internalising symptoms that would be overestimated if simple differences were calculated. The simplest empirical set up of the fixed effects difference-in-differences design can be expressed by the formula below.

$$\boldsymbol{Y_i g_t = \beta_0 + \beta_1 T_g + \beta_2 P_t\, t + \beta_3\, (T_g \times P_t) + f_{it} + g_t}$$

Equation 6. 1

Where:

$\boldsymbol{Y_i g_t}$ =average parent- or teacher-reported externalising or internalising difficulties for the i[th] child during study period t

$\boldsymbol{\beta_0}$ = Baseline average

$\boldsymbol{\beta_1 T_g}$ = Time trend in the control group

$\boldsymbol{\beta_2 P_t\, t}$ =Difference between control and treatment group before the intervention.

$\boldsymbol{\beta_3\, (T_g \times P_t)}$ =The difference in changes over time.

$\boldsymbol{f_{it}}$ is the child fixed effect for for the i[th] child capturing time invariant child characteristics during study period t.

$\boldsymbol{g_t}$ = Time-varying error

It is also likely that the difference-in-difference might vary as result of a third variable δ (Wing et al., 2018, Afendulis et al., 2011, Chetty et al., 2009). In this case variable δ represents either family income or verbal cognitive skills at MCS 3. The *equation 6.1* was then extended to include $T_g \times P_t \times \delta$ forming a triple differences model.

As was mentioned in chapter 4, caution should be exercised when interpreting three-way interactions since these interactions could be underpowered (Heo and Leon, 2010). Besides, the hierarchical structure of the data used to perform difference-in-differences regression modelling in this chapter could address effectively the issue of multiple comparisons and subsequent false discoveries, and produce more efficient estimates (Gelman et al., 2012).

6.5: Model assumptions

The parallel growths assumption (Reggio and Mora, 2012) or difference in the first differences is the main identifying assumption. As was mentioned in chapter 5, parallel growths assumption highlights that in the absence of the transition to attainment grouping, average changes in outcome among those who were grouped by attainment at age 11 will be equal to the average changes among those MCS participants who were never grouped by attainment. A deviation from the trend line of the control group identifies a change in the treatment effect.

Another assumption of the difference-in-differences is that there is no time-varying confounding associated with only the transition to attainment grouping and average changes in children's externalising or internalising symptoms.

6.6: Results

6.6.1: Distribution of children's externalising and internalising symptoms at MCS 4 and MCS 5.

The distributions of parent-reported and teacher-reported externalising and internalising psychological symptoms at ages 7 and 11 years were skewed to the right for both variables with higher percentages of lower scores at each age (see Figures 6.1 to 6.4). Interestingly, the distribution of parent-reported externalising symptoms at both ages 7 and 11 years was more symmetric than the distribution of teacher-reported symptoms at both ages. A higher proportion of zeros in teacher responses revealed important differences on how parents and teachers perceive and report children's externalising psychological symptoms. The distributions of parent-reported and teacher-reported internalising symptoms were similar. A higher proportion of internalising symptoms was reported by teachers than parents, signalling again differences in the perception and report of internalising symptoms between parents and teachers. The distributions of both parent- and teacher-reported externalising and internalising psychological symptoms at ages 7 and 11 were not normal. Nevertheless, this was not a problem for the validity of estimates due to the large enough sample (N>500) of MCS (Lumley et al., 2002).

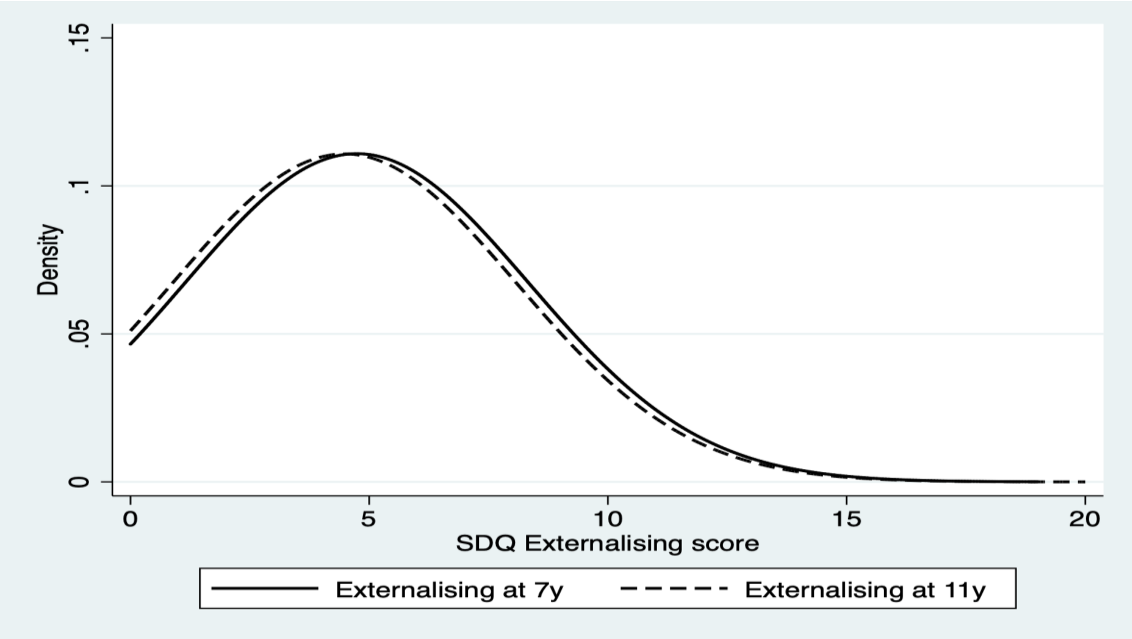

Figure 6. 1: Distribution parent-reported externalising symptoms s at age 7 and 11 years (N=7,145)

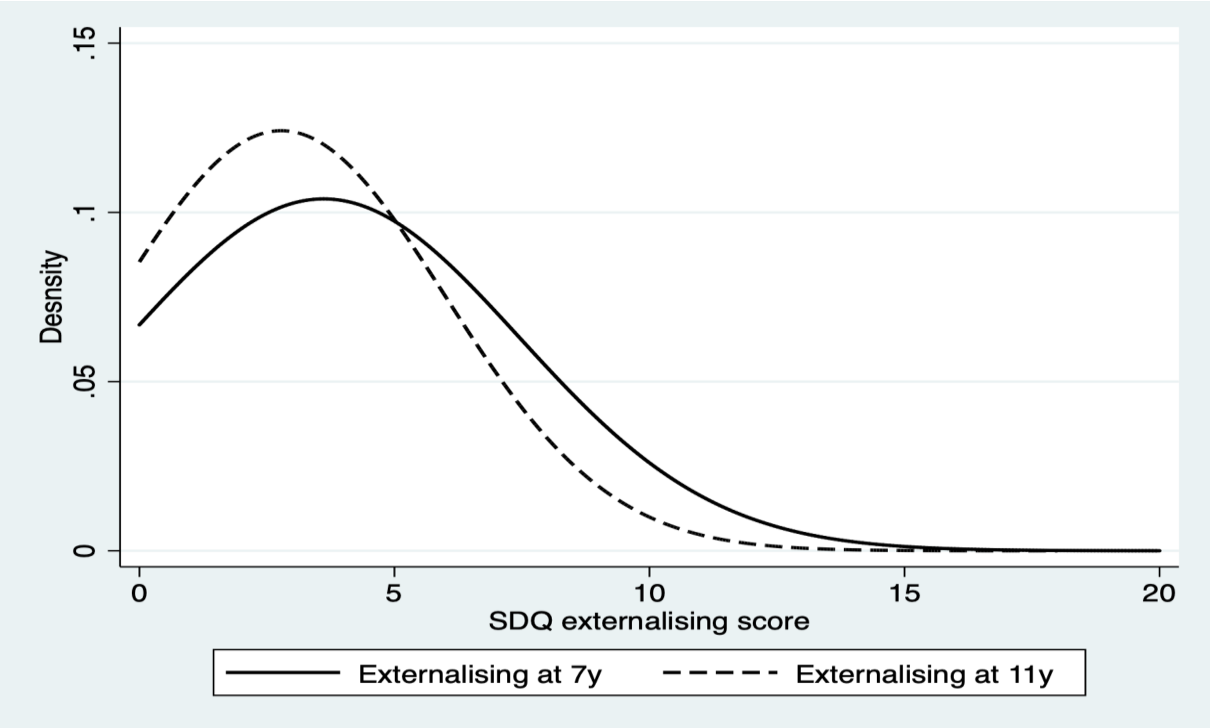

Figure 6. 2: Distribution of teacher-reported externalising symptoms at age 7 and 11 years (N=,7169)

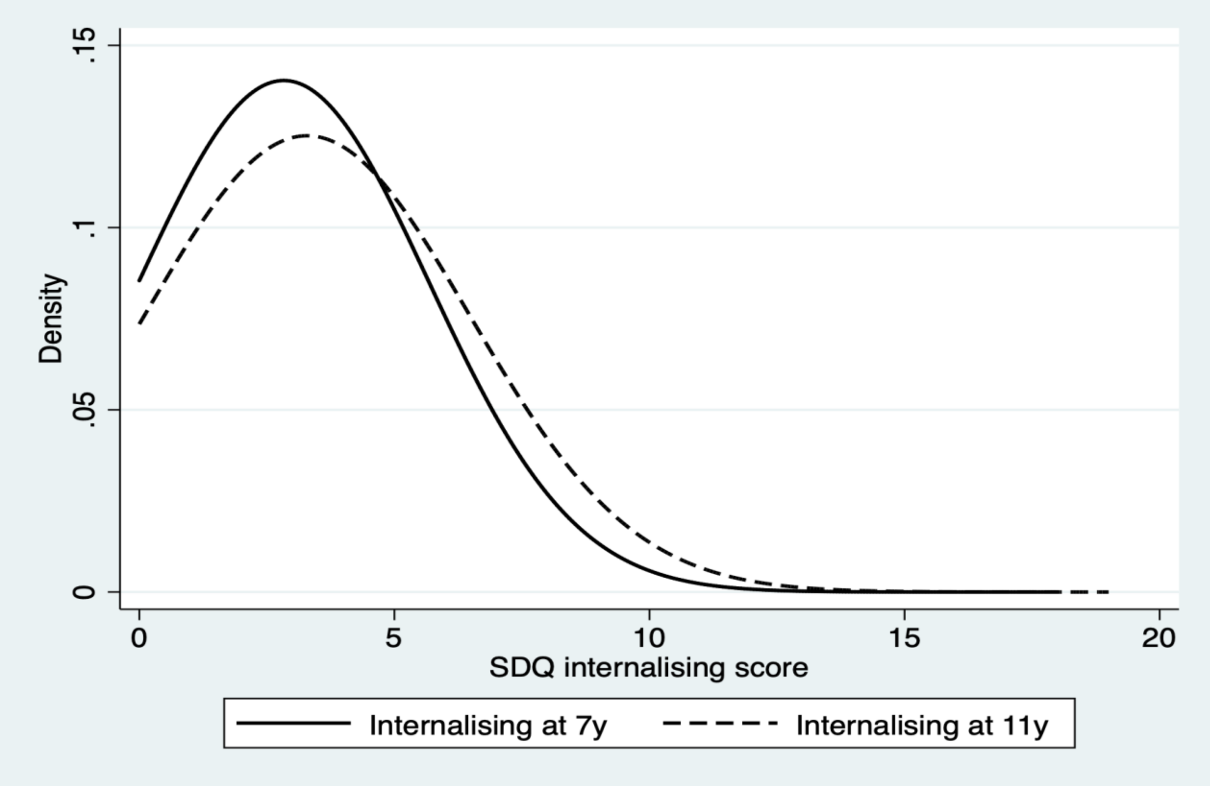

Figure 6. 3: Distribution of parent-reported internalising symptoms at age 7 and 11 years (N=7,145)

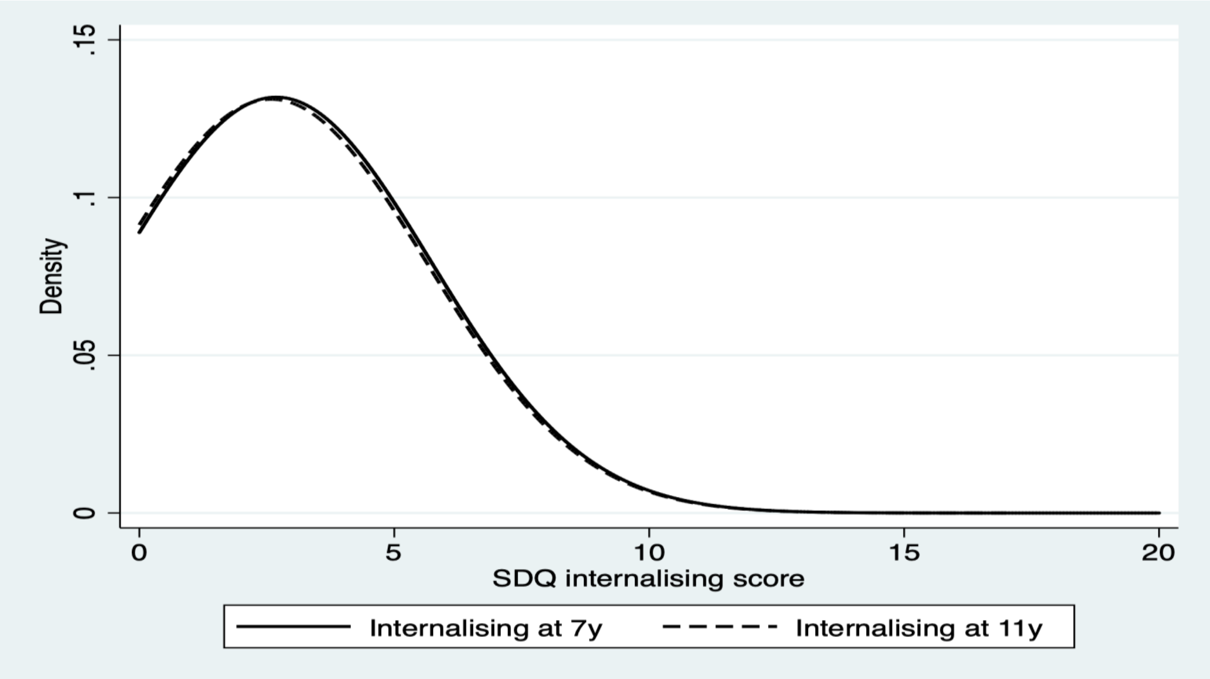

Figure 6. 4: Distribution of teacher-reported internalising symptoms s at age 7 and 11 (N=7,169)

Children’s mean externalising symptoms as reported by the MCS participant’s main carer and teacher significantly decreased from age 7 to age 11 (see Figure 6.5). This decrease was more pronounced for teacher-reported externalising symptoms. Children’s mean internalising symptoms as reported by MCS participant’s main carer significantly increased from age 7 to age 11. In contrast, children’s teacher-reported mean symptoms of internalising difficulties did not significantly change from age 7 to age 11 (see Figure 6.6).

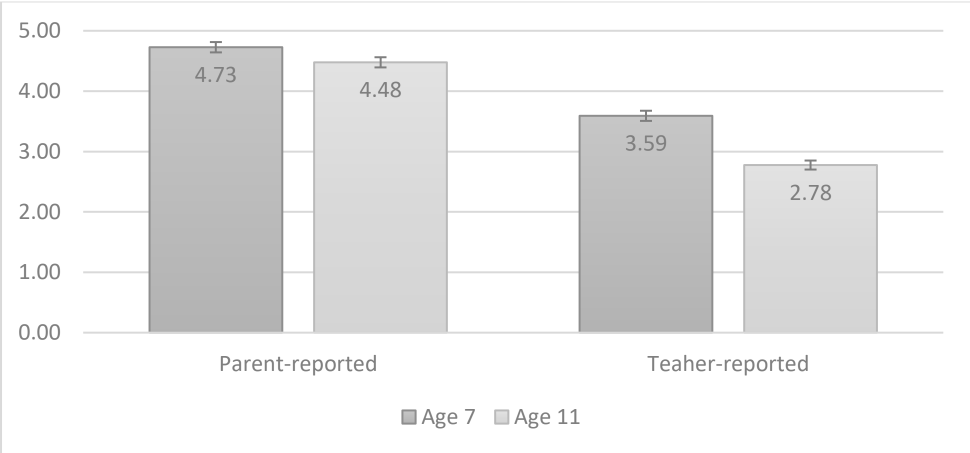

Figure 6. 5: Mean parent- and teacher-reported externalising symptoms at age 7 and age 11 years

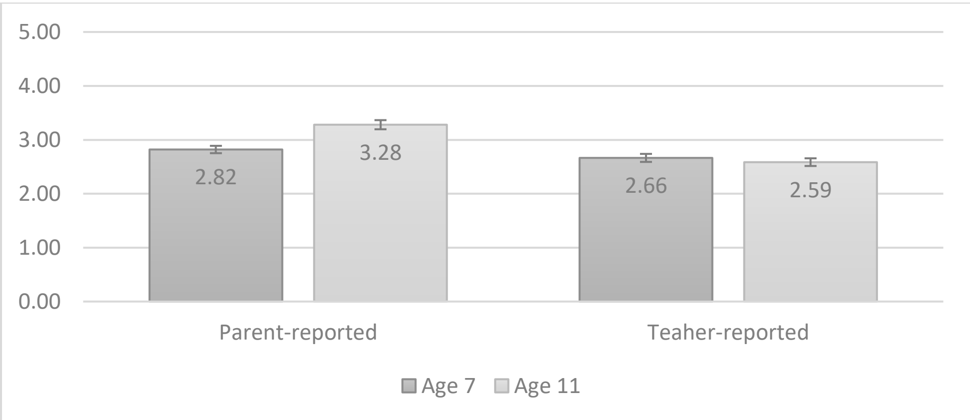

Figure 6. 6: Mean parent- and teacher-reported internalising symptoms at age 7 and age 11 years

6.7: Difference-in Difference analysis

6.7.1: Transition to attainment grouping and children's externalising or internalising symptoms from age 7 to age 11

Tables 6.2 and 6.3 and Figures 6.9 and 6.10 show a significant decrease in both parent-reported and teacher-reported externalising symptoms between ages 7 and 11years using both complete cases (-0.25 and -0.69 respectively) and multiple imputed data (-0.26 and -0.62 respectively). There was also a difference in how parents and teachers reported externalising difficulties over this four-year period, with teachers reporting a significantly bigger decrease in externalising difficulties (-0.62 points in the SDQ externalising subscale) compared to parents (-0.26 points in the SDQ externalising subscale). No evidence of a statistically significant association was observed between the transition to attainment grouping and changes in both parent- and teacher-reported symptoms of externalising symptoms from age 7 to age 11.

Table 6.4 and Figure 6.11 highlight a significant increase in parent-reported internalising symptoms from age 7 to age 11 by using both complete cases and multiple imputed data. A small, but non-statistically significant increase by 0.14 points was observed in teacher-reported internalising symptoms from age 7 to age 11 was observed. There was also no evidence of an association between the change in both parent-reported and teacher-reported internalising symptoms, and attainment grouping transition between ages 7 and 11 years.

Table 6. 2: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11

	Complete case analysis(n=2,031)				**Multiple imputation (n=4,157)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-0.25*	0.10	-0.44	-0.06	-0.26**	0.09	-0.44	-0.07
Attainment grouping * MCS 5	0.05	0.12	-0.19	0.29	0.06	0.13	-0.19	0.31

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

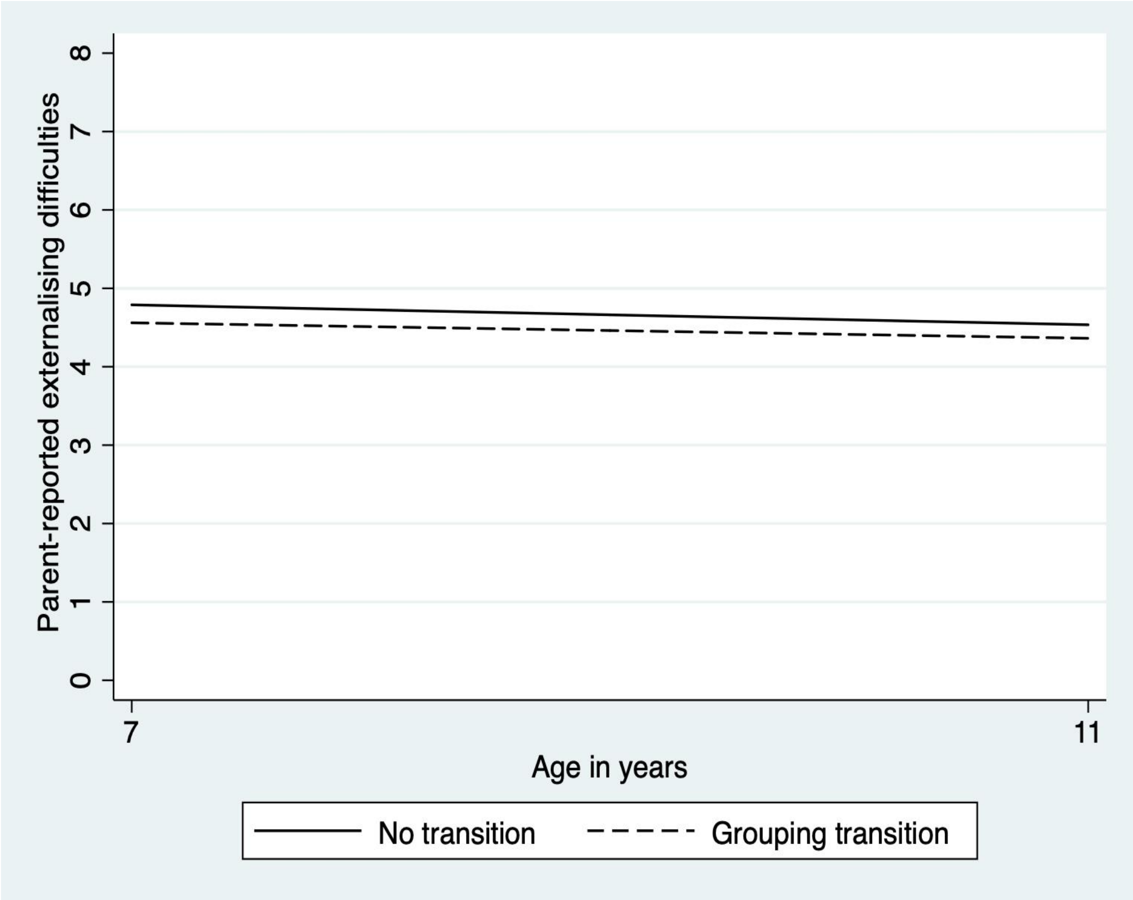

Figure 6. 7: *Transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11*

Table 6. 3: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported externalising symptoms form age 7 to age 11

	Complete case analysis (n=2,190)				**Multiple imputation (n=4,181)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-0.69***	0.11	-0.90	-0.48	-0.62***	0.10	-0.82	-0.41
Attainment grouping * MCS 5	-0.10	0.14	-0.37	0.16	-0.12	0.13	-0.39	0.14

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

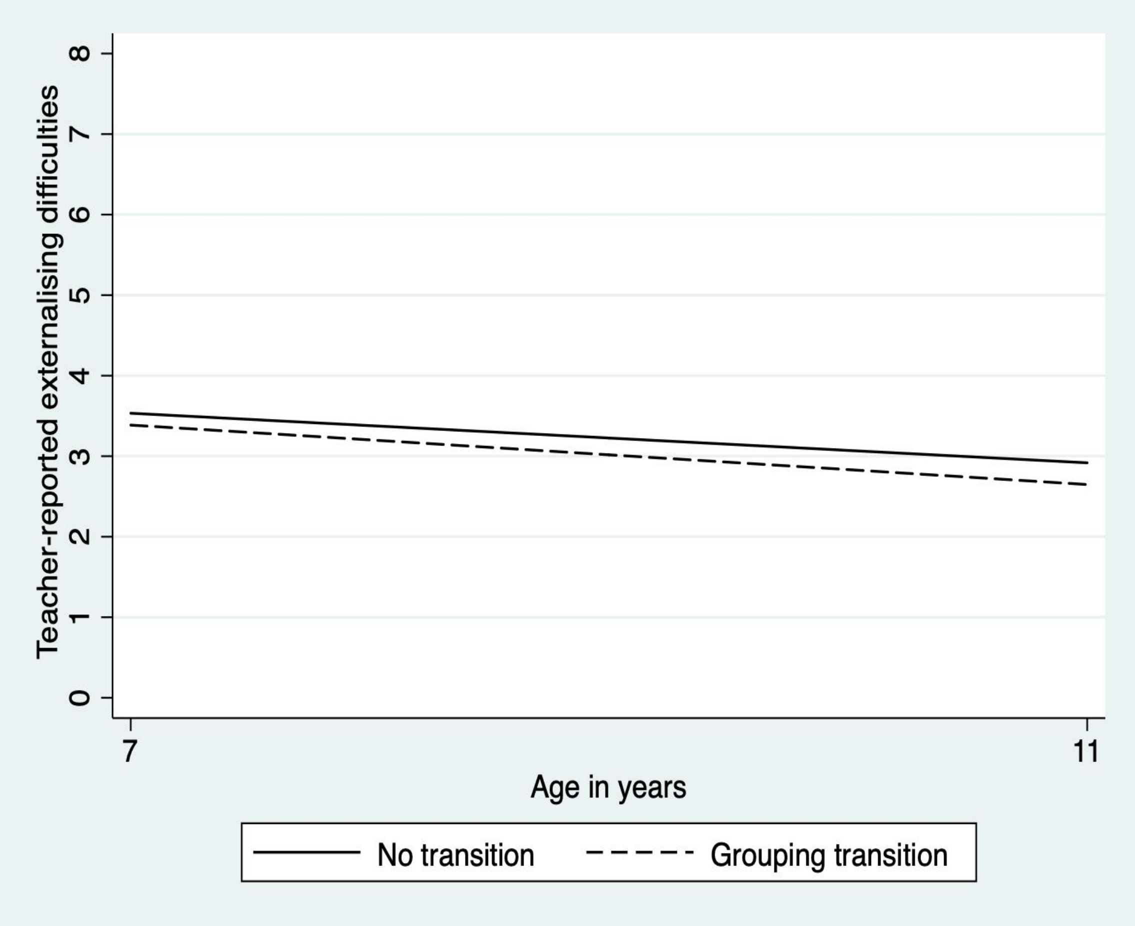

Figure 6. 8: *Transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported externalising symptoms from age 7 to age 11 (n=4,181).*

Table 6. 4: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported internalising symptoms from age 7 to age 11

	Complete case analysis (n=2,031)				**Multiple imputation (n=4,157)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	0.61***	0.09	0.43	0.79	0.53***	0.10	0.33	0.74
Attainment grouping * MCS5	-0.19	0.12	-0.42	0.04	-0.11	0.13	-0.37	0.15

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

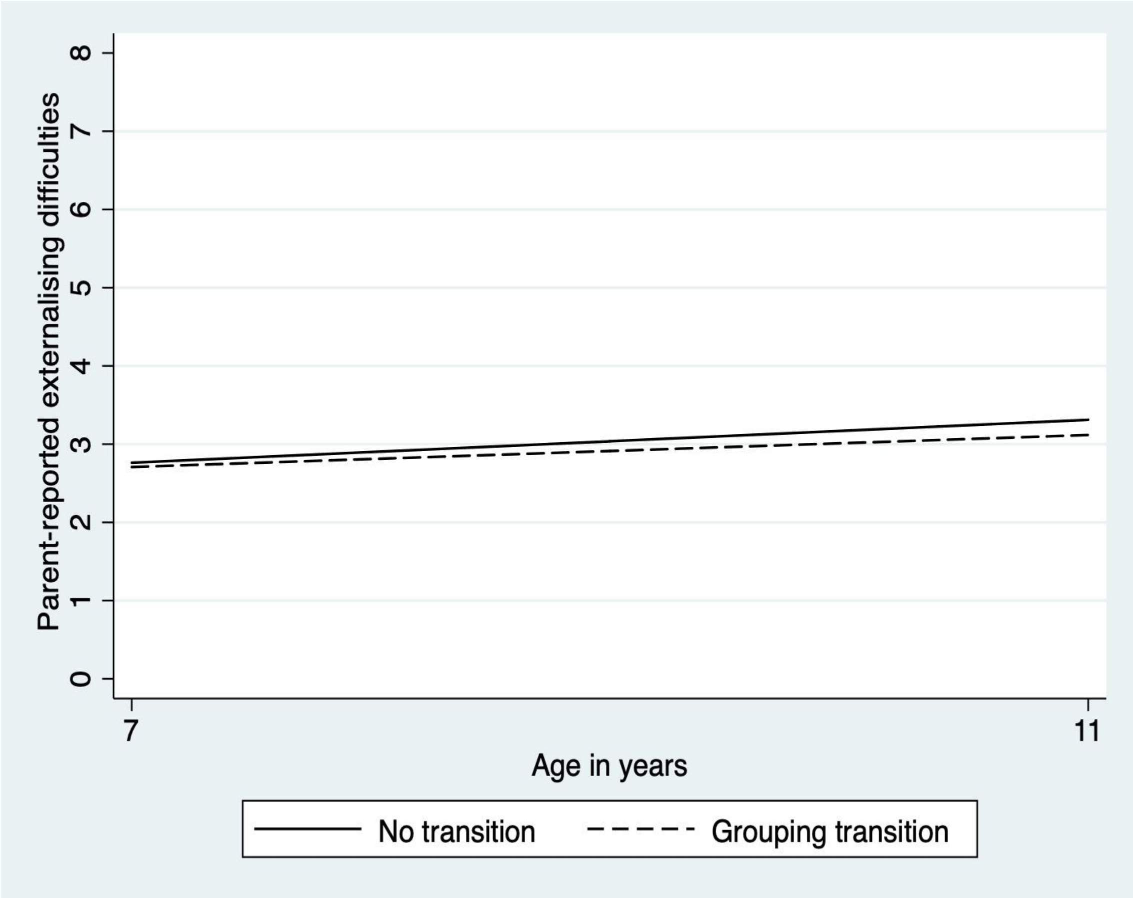

Figure 6. 9: Transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported internalising symptoms from age 7 to age 11 (n=4,157).

Table 6. 5: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported internalising symptoms from age 7 to age 11

	Complete case analysis (n=2,190)				**Multiple imputation (n=4,181)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	0.08	0.12	-0.15	0.30	0.14	0.11	-0.07	0.35
Attainment grouping *MCS5	-0.05	0.15	-0.34	0.23	-0.06	0.14	-0.34	0.22

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

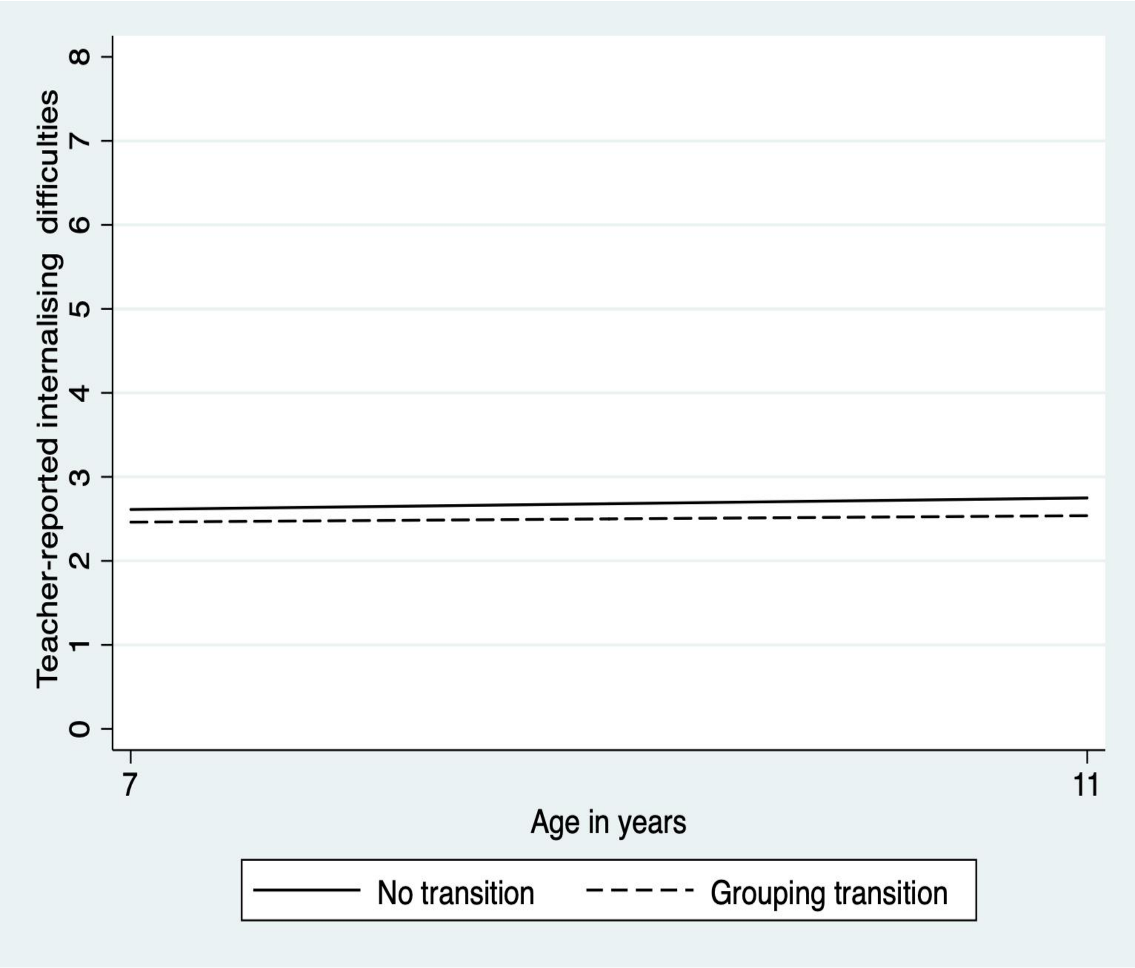

Figure 6. 10: *Transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported internalising symptoms from age 7 to age 11*

6.7.2: Transition to attainment grouping and children's externalising or internalising symptoms from age 7 to age 11

The findings obtained from the fixed effects difference-in-difference-in-differences modelling, which shown in Tables 6.6 to 6.9 and Figures 6.13 to 6.16, are suggestive of no effect modification by family income. This means that children who changed to attainment grouping between ages 7 and 11 years experienced similar changes in parent-reported and teacher-reported externalising and internalising from age 7 to age 11 as children who did not change to attainment grouping.

Table 6. 6: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11 *by family income*

	Complete case analysis(n=2,031)				Multiple imputation (n=4,157)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-0.26	0.18	-0.62	0.10	-0.30	0.15	-0.60	0.00
Attainment grouping * MCS5	0.10	0.22	-0.33	0.52	0.07	0.19	-0.31	0.45
Attainment grouping * Highest incomes (Ref.)								
2nd quantile * MCS5	-0.05	0.25	-0.53	0.43	-0.04	0.20	-0.44	0.36
3rd quantile * MCS5	-0.14	0.26	-0.65	0.37	0.00	0.23	-0.45	0.45
4th quantile * MCS5	-0.02	0.35	-0.70	0.66	0.01	0.26	-0.50	0.51
Lowest incomes * MCS5	0.52	0.40	-0.25	1.30	0.31	0.31	-0.30	0.92
Attainment grouping x Highest incomes * MCS5 (Ref.)								
Attainment grouping * 2nd quantile * MCS5	0.16	0.30	-0.43	0.75	0.14	0.27	-0.39	0.67
Attainment grouping * 3rd quantile * MCS5	0.07	0.34	-0.60	0.73	0.01	0.31	-0.61	0.63
Attainment grouping * 4th quantile * MCS5	-0.14	0.42	-0.96	0.67	0.02	0.33	-0.64	0.68
Attainment grouping * Lowest incomes * MCS5	-0.76	0.52	-1.79	0.27	-0.28	0.43	-1.13	0.56

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

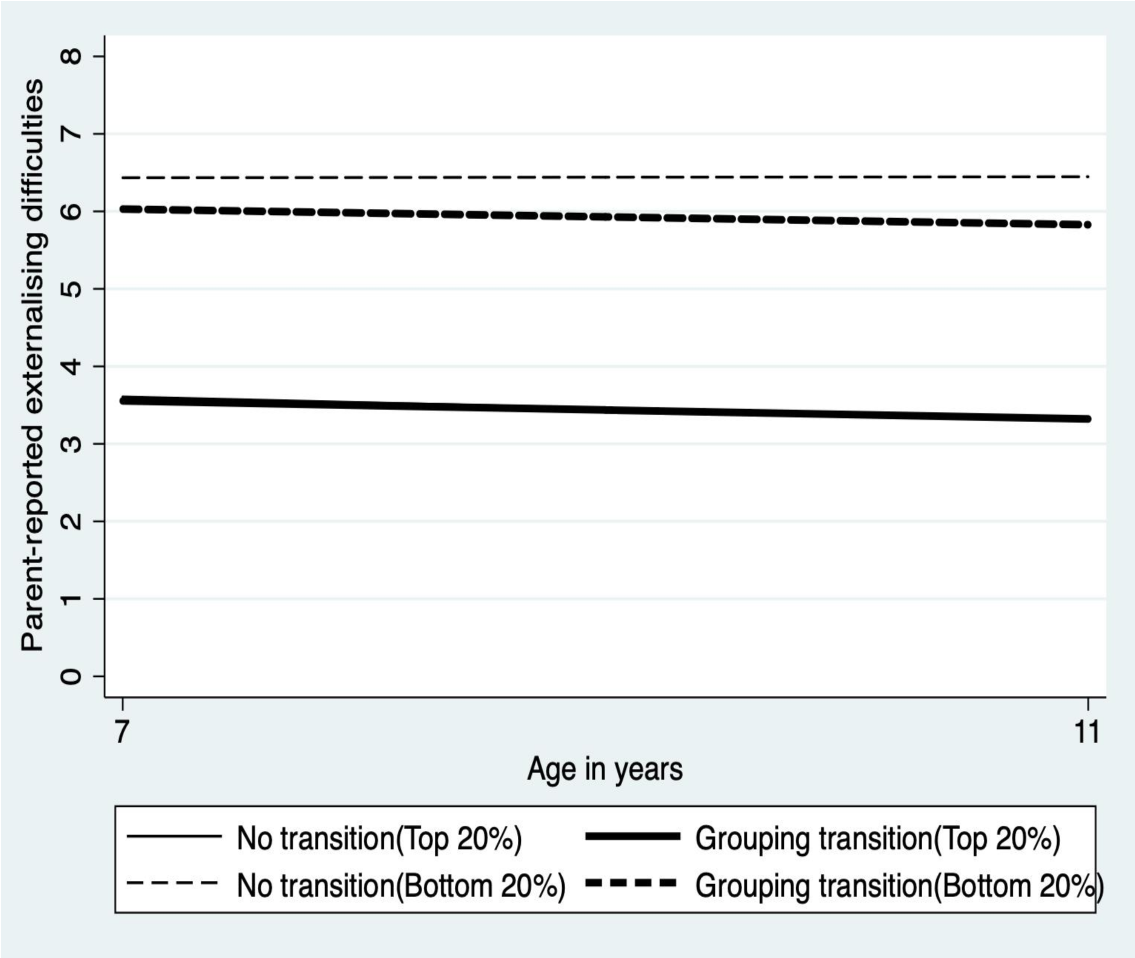

Figure 6. 11: Transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11 by family income (n=4,157)

Table 6. 7: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported externalising symptoms from age 7 to age 11 by family income

	Complete case analysis (n=2,031)				**Multiple imputation (n=4,157)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 and MCS 5	-0.48*	0.21	-0.89	-0.06	-0.54**	0.18	-0.90	-0.19
Attainment grouping * MCS5	-0.25	0.26	-0.75	0.26	-0.21	0.23	-0.66	0.24
Attainment grouping * Highest incomes (Ref.)								
2nd quantile * MCS5	-0.51	0.28	-1.05	0.03	-0.30	0.26	-0.81	0.21
3rd quantile * MCS5	-0.52	0.32	-1.15	0.10	-0.31	0.27	-0.83	0.22
4th quantile * MCS5	0.43	0.39	-0.33	1.19	0.34	0.31	-0.27	0.95
Lowest incomes * MCS5	-0.16	0.38	-0.90	0.57	0.19	0.33	-0.46	0.84
Attainment grouping * Highest incomes * MCS5 (Ref.)								
Attainment grouping * 2nd quantile * MCS5	0.61	0.35	-0.07	1.30	0.43	0.33	-0.23	1.09
Attainment grouping * 3rd quantile * MCS5	0.32	0.40	-0.46	1.10	0.19	0.35	-0.50	0.88
Attainment grouping * 4thquantile * MCS5	-0.80	0.49	-1.75	0.15	-0.44	0.40	-1.22	0.35
Attainment grouping * Lowest incomes * MCS5	0.21	0.49	-0.75	1.17	0.07	0.43	-0.77	0.91

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

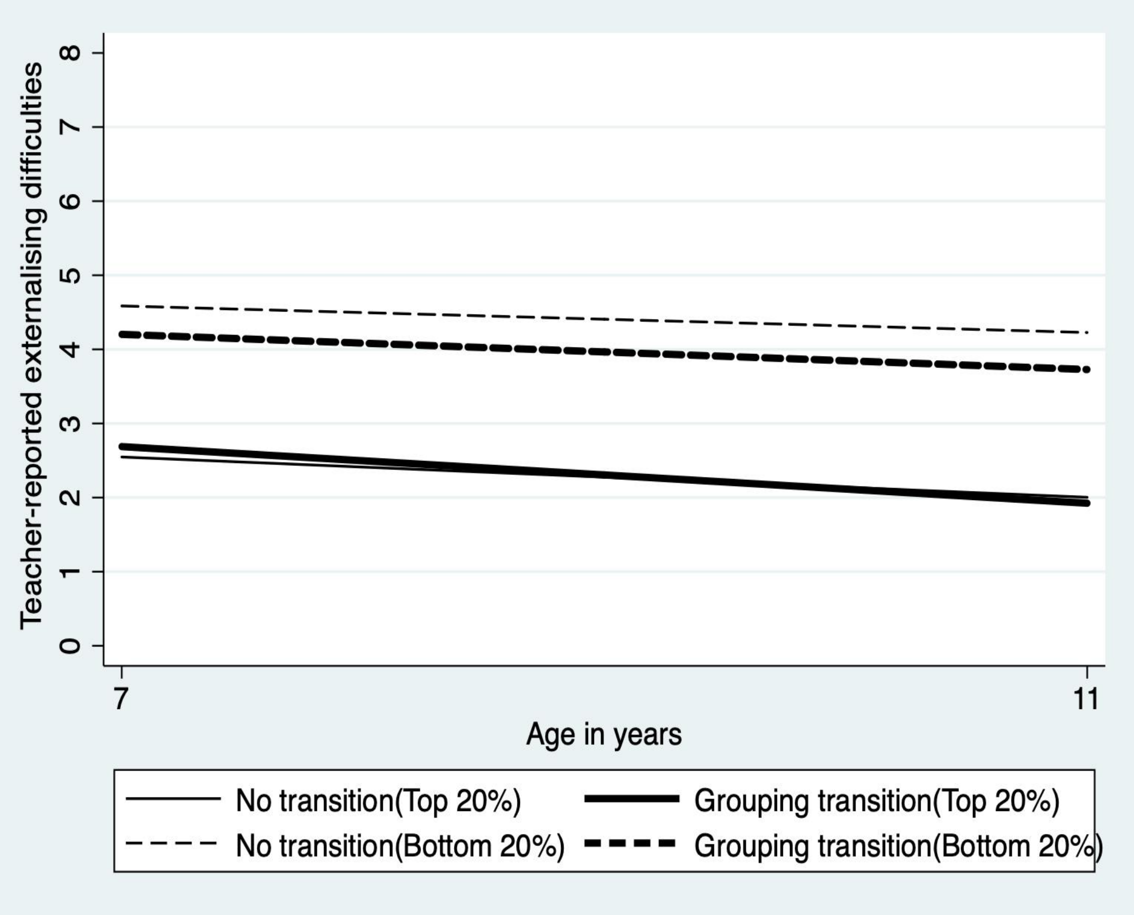

Figure 6. 12: Transition to attainment grouping between ages 7 and 11y ears and changes in children's teacher-reported externalising difficulties from age 7 to age 11 by family income (n=4,157)

Table 6. 8: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported internalising symptoms from age 7 to age 11 by family income

	Complete case analysis (n=2,031)				**Multiple imputation (n=4,157)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	0.32*	0.16	0.01	0.64	0.39*	0.16	0.08	0.70
Attainment grouping *MCS5	-0.22	0.20	-0.62	0.18	0.07	0.20	-0.32	0.45
Highest incomes (Top quantile) * MCS5 (Ref.)								
2nd quantile * MCS5	0.34	0.23	-0.10	0.79	0.20	0.21	-0.22	0.63
3rd quantile * MCS5	0.68	0.26	0.17	1.18	0.41	0.23	-0.04	0.86
4th quantile * MCS5	0.19	0.31	-0.42	0.79	0.18	0.27	-0.35	0.71
Lowest incomes * MCS5	0.12	0.36	-0.59	0.82	-0.08	0.230	-0.66	0.50

	Complete case analysis (n=2,031)				Multiple imputation (n=4,157)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Attainment grouping *Highest incomes * MCS5 (Ref.)								
Attainment grouping * 2nd quantile * MCS5	0.10	0.29	-0.48	0.67	0.07	0.28	-0.47	0.61
Attainment grouping * 3rd quantile * MCS5	-0.25	0.33	-0.89	0.39	-0.06	0.30	-0.66	0.54
Attainment grouping * 4thquantile * MCS5	0.31	0.38	-0.44	1.07	0.15	0.35	-0.53	0.83
Attainment grouping * Lowest incomes * MCS5	0.19	0.45	-0.71	1.09	0.05	0.40	-0.74	0.84

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

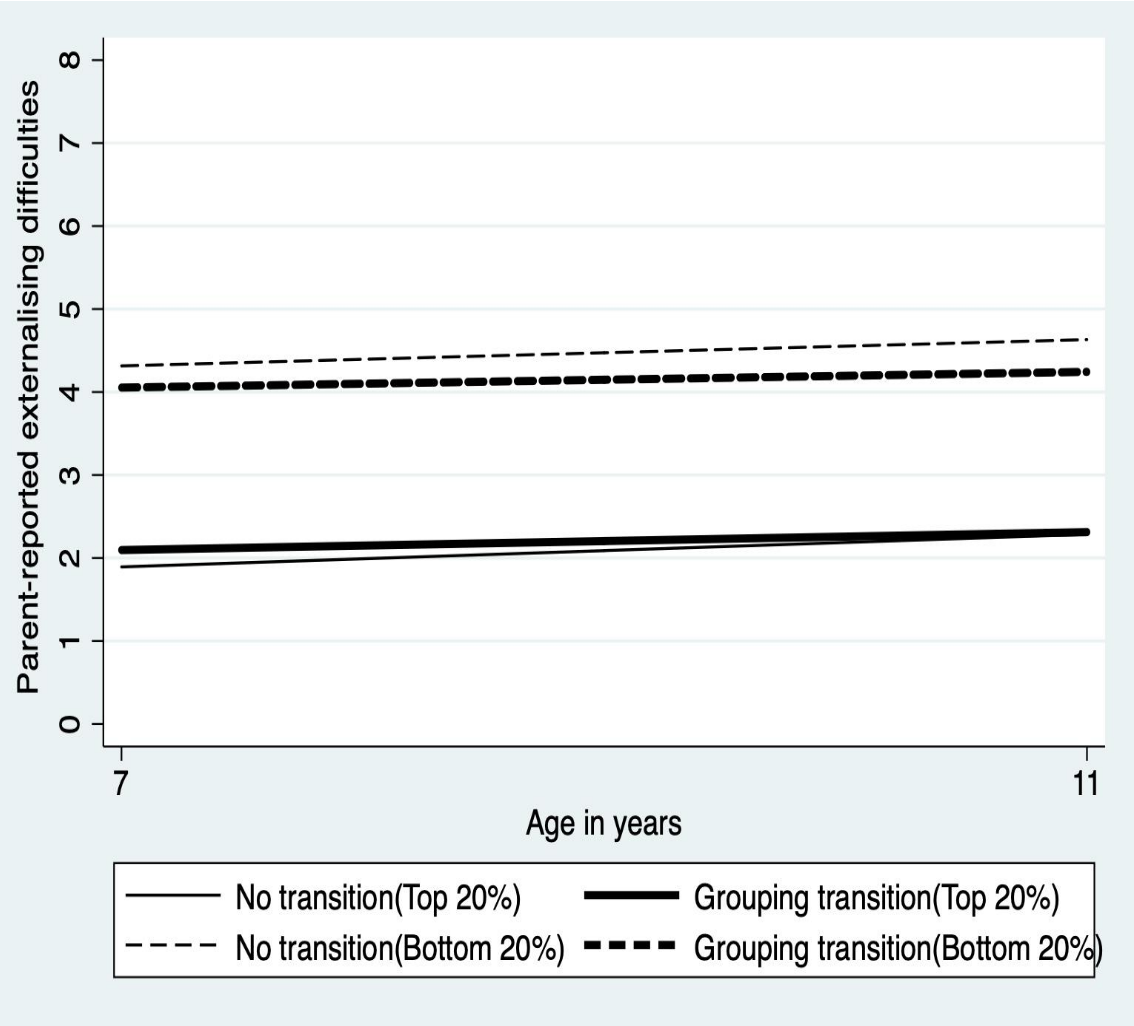

Figure 6. 13: Transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported internalising symptoms from age 7 to age 11 by family income (n=4,157).

Table 6. 9: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported internalising symptoms from age 7 to age 11 by family income

	Complete case analysis (n=2,190)				**Multiple imputation (n=4,181)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-0.26	0.18	-0.61	0.09	-0.16	0.19	-0.54	0.22
Attainment grouping*MCS5	0.09	0.23	-0.36	0.54	0.06	0.25	-0.43	0.54
Highest incomes * MCS5 (Ref.)								
2nd quantile * MCS5	0.24	0.28	-0.31	0.78	0.21	0.26	-0.31	0.72
3rd quantile * MCS5	0.36	0.30	-0.23	0.94	0.24	0.28	-0.32	0.80
4th quantile * MCS5	0.90	0.41	0.10	1.71	0.58	0.36	-0.12	1.29
Lowest incomes * MCS5	0.43	0.42	-0.39	1.26	0.57	0.33	-0.08	1.23

	Complete case analysis (n=2,190)				Multiple imputation (n=4,181)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Attainment grouping * Highest incomes *MCS5 (Ref.)								
Attainment grouping * 2nd quantile * MCS5	0.14	0.36	-0.56	0.84	0.04	0.33	-0.60	0.69
Attainment grouping * 3rd quantile * MCS5	-0.26	0.38	-1.00	0.47	-0.13	0.38	-0.87	0.61
Attainment grouping * 4th quantile * MCS5	-0.83	0.51	-1.83	0.16	-0.43	0.46	-1.33	0.48
Attainment grouping * Lowest incomes * MCS5	0.13	0.55	-0.94	1.21	-0.07	0.45	-0.96	0.81

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

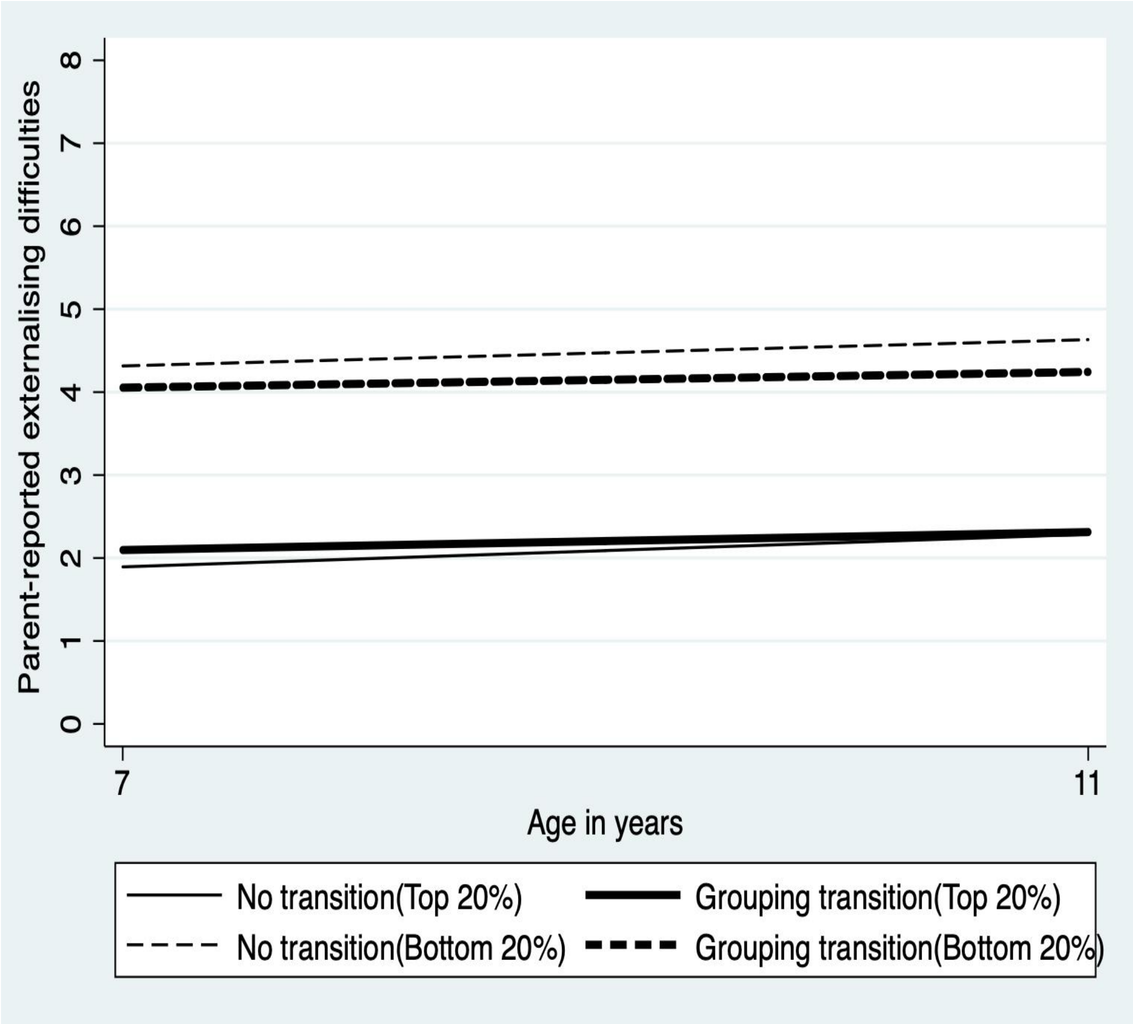

Figure 6. 14: Transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported internalising symptoms from age 7 to age 11 by family income (n=4,181)

6.7.3: Transition to attainment grouping and children's externalising and internalising symptoms by baseline verbal skills

The findings obtained from the fixed effects difference-in-difference-in-differences modelling and shown in Tables 6.10 to 6.13 and Figures 6.17 to 6.20 were suggestive of effect modification by initial verbal cognitive skills only for parent-reported externalising symptoms and using complete cases. Children who had higher levels of verbal skills in the beginning of primary school in England and Wales and who subsequently changed to attainment grouping experienced a slightly smaller decrease in their parent-reported externalising difficulties between 7 and 11 years compared to children who started primary school with a higher level of verbal skills and who did not change to attainment grouping (see Table 6.10). Nonetheless, this interaction coefficient between attainment grouping transition, time and baseline verbal skills did not remain significant at the 5% level when multiple imputed data were used.

Table 6. 10: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11 by baseline verbal skills

	Complete cases (N=2,031)				**Multiple imputation (N=4,157)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	0.86	0.55	-0.22	1.94	0.36	0.48	-0.58	1.30
Attainment grouping * MCS5	-1.38*	0.68	-2.71	-0.05	-0.84	0.60	-2.01	0.33
Baseline verbal cognitive skills * MCS5	-0.02*	0.01	-0.04	0.00	-0.01	0.01	-0.03	0.01
Baseline verbal cognitive skills * attainment grouping * MCS5	0.03*	0.01	0.00	0.05	0.02	0.01	0.00	0.04

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

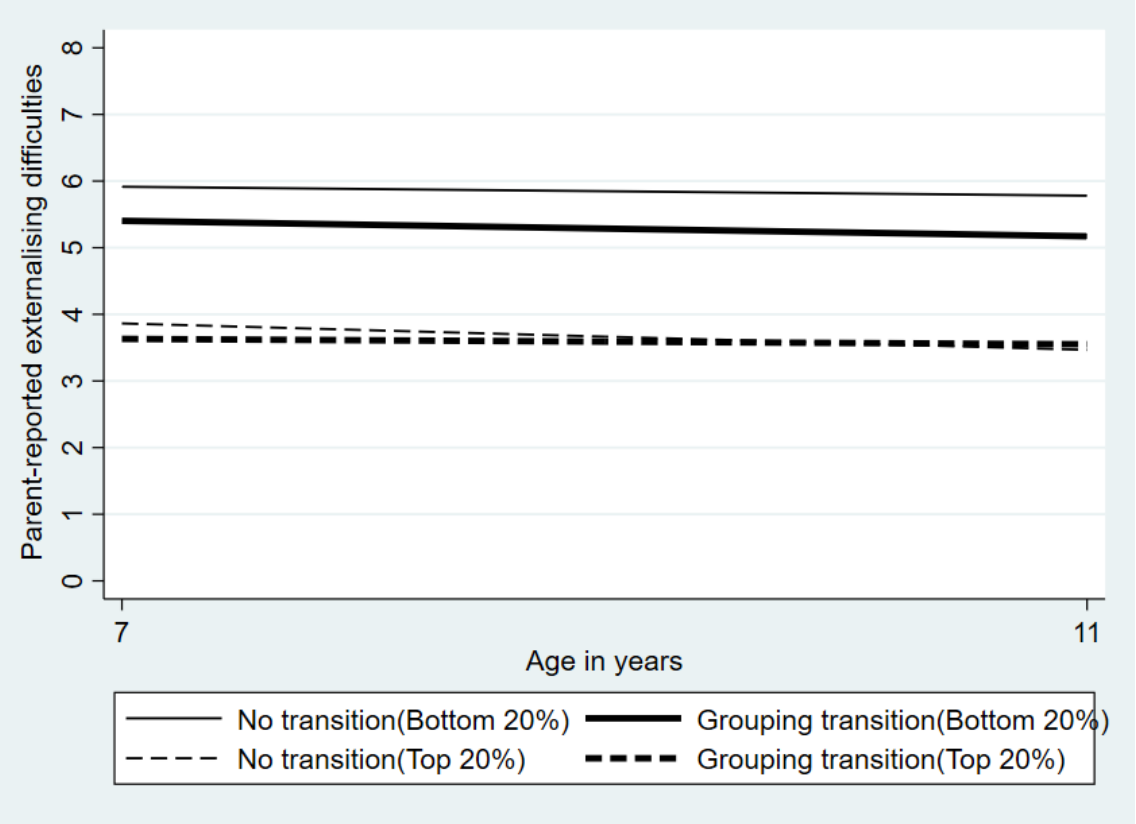

Figure 6. 15: *Attainment grouping between ages 7 and 11 years and changes in children's parent-reported externalising symptoms from age 7 to age 11 by baseline verbal skills (n=4,157)*

Table 6. 11: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported externalising symptoms from age 7 to age 11 by baseline verbal skills

	Complete cases (N=2,190)				**Multiple imputation (N=4,181)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-1.20**	0.59	-2.35	-0.05	-0.99*	0.50	-1.96	-0.01
Attainment grouping * MCS5	-0.37	0.74	-1.83	1.09	-0.38	0.66	-1.69	0.92
Baseline verbal cognitive skills * MCS5	0.01	0.01	-0.01	0.03	0.01	0.01	-0.01	0.02
Baseline verbal cognitive skills * attainment grouping * MCS5	0.01	0.01	-0.02	0.03	0.00	0.01	-0.02	0.03

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

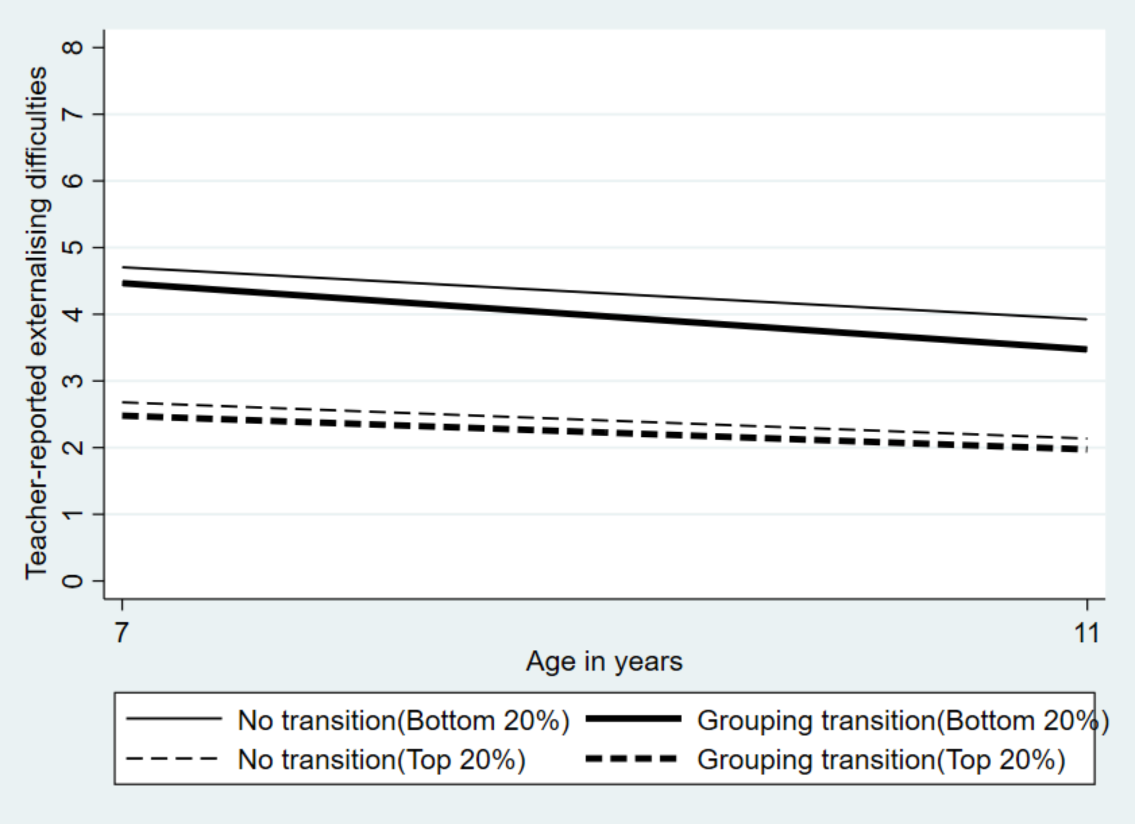

Figure 6. 16: Attainment grouping between ages 7 and 11 years and changes in children's teacher-reported externalising symptoms from age 7 to age 11 by baseline verbal skills (n=4,181).

Table 6. 12: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's parent-reported internalising symptoms from age 7 to age 11 by baseline verbal skills

	Complete cases (N=2,031)				Multiple imputation (N=4,157)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.25*	0.56	0.15	2.35	0.97	0.49	0.00	1.94
Attainment grouping * MCS5	-0.03	0.70	-1.40	1.35	-0.31	0.64	-1.57	0.95
Baseline verbal cognitive skills * MCS5	-0.01	0.01	-0.03	0.01	-0.01	0.01	-0.02	0.01
Baseline verbal cognitive skills *attainment grouping * MCS5	0.00	0.01	-0.03	0.02	0.00	0.01	-0.02	0.02

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

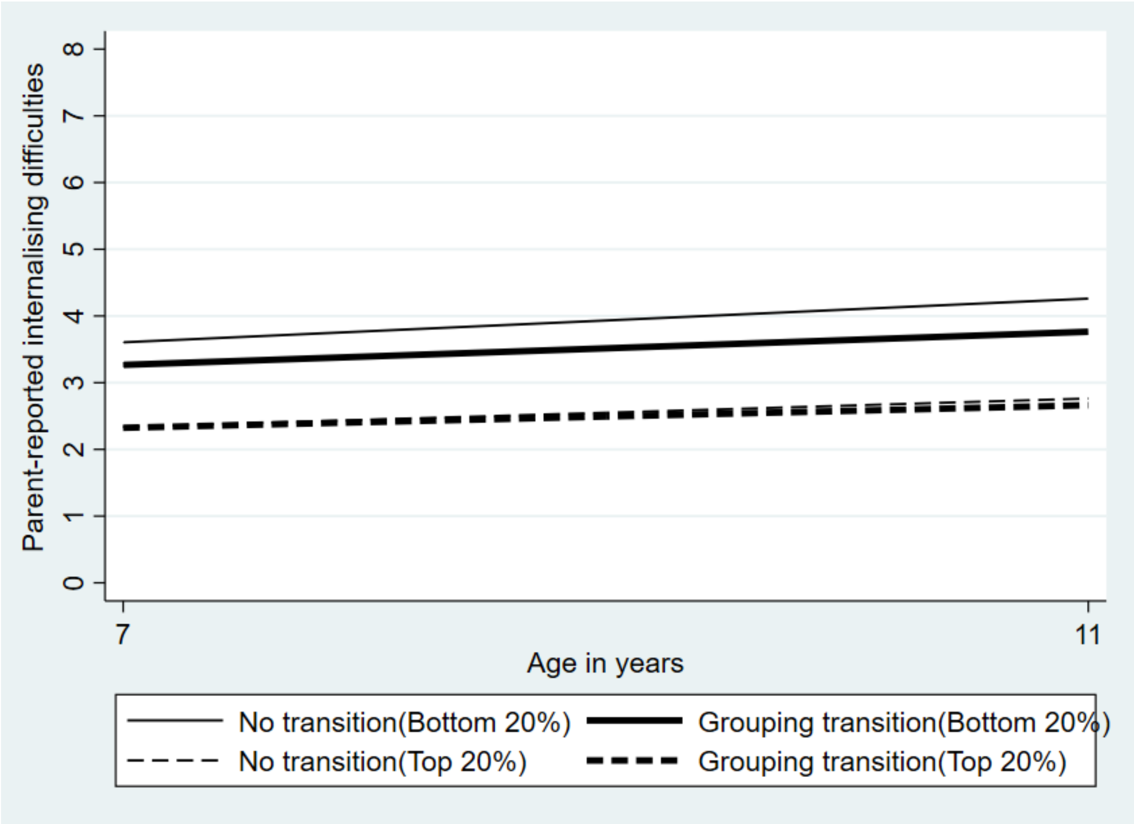

Figure 6. 17: Attainment grouping between ages 7 and 11 years and changes in children's *parent*-reported *internalising symptoms* from age 7 to age 11 by baseline verbal skills (n=4,157)

Table 6. 13: The association between transition to attainment grouping between ages 7 and 11 years and changes in children's teacher-reported internalising symptoms from age 7 to age 11 by baseline verbal skills

	Complete cases (N=2,190)				**Multiple imputation (N=4,181)**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	
Change from MCS 4 to MCS 5	-0.10	0.69	-1.45	1.25	0.22	0.60	-0.95	1.40
Attainment grouping * MCS5	0.60	0.85	-1.06	2.27	0.28	0.74	-1.18	1.74
Baseline verbal cognitive skills * MCS5	0.00	0.01	-0.02	0.03	0.00	0.01	-0.02	0.02
Baseline verbal cognitive skills *attainment grouping *MCS5	-0.01	0.01	-0.04	0.02	-0.01	0.01	-0.03	0.02

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

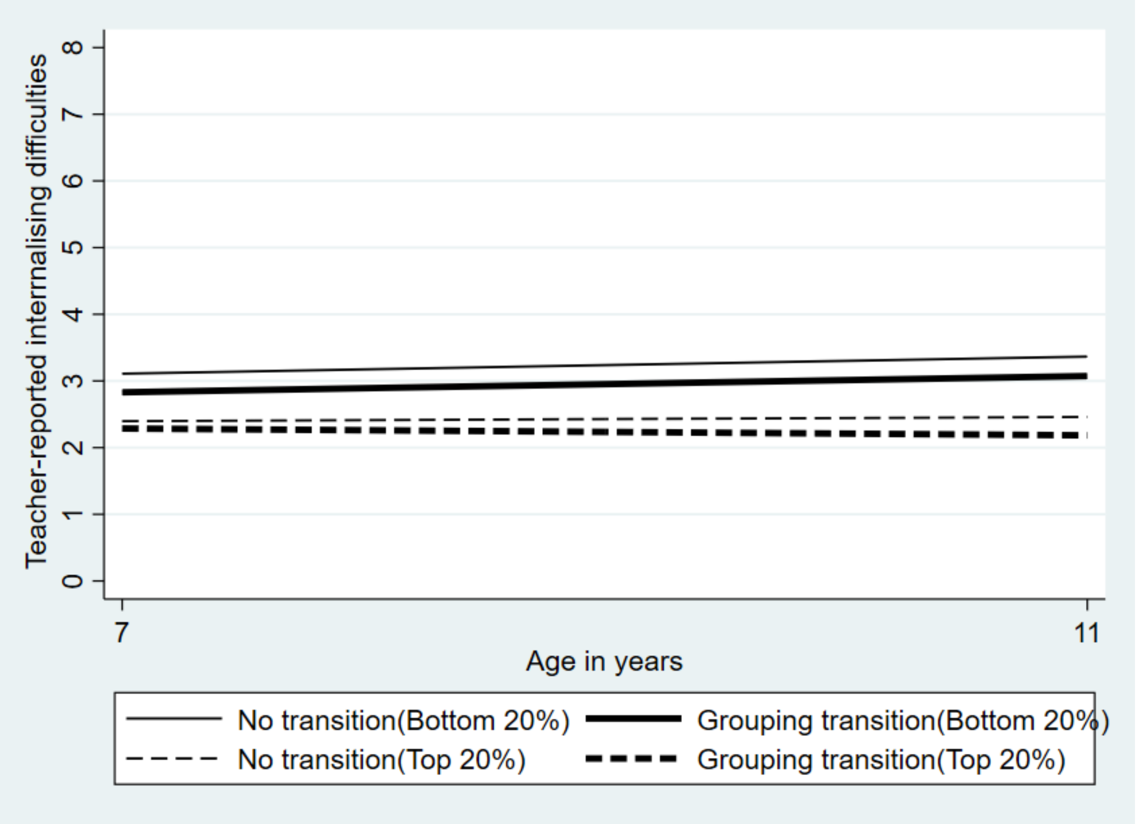

Figure 6. 18: Attainment grouping between ages 7 and 11 years and changes in children's *parent*-reported *internalising symptoms from age 7 to age 11* by baseline l verbal skills (n=4,181).

6.7.4: Sensitivity analysis

Sensitivity analysis was conducted using complete cases since multiple imputation data in this PhD yield similar estimates of similar statistical significance.

6.7.4.1: The influence of attainment grouping placement on inequalities in children's psychological symptoms between MCS 4 and MCS 5

The influence of attainment grouping transition on the change in children's psychological symptoms for those with different family socioeconomic circumstances and baseline verbal skills (MCS 3-age 5) may vary according to placement to the top, middle or bottom streams or sets. To explore this possibility, stream and set group placements were added as additional controls.

Tables A3.6 and A3.7 in Appendix 3 show that there was little evidence of an association between transition to attainment grouping and change in psychological symptoms among children who changed to attainment grouping between age 7 and age 11 after adjusting for set or stream placement

6.8: Summary of findings

What follows is a summary of the findings of this chapter's analyses addressing objectives 2.1 to 2.3 and their related hypotheses.

Objective 2.1: To examine whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a less favourable change in children's externalising or internalising psychological symptoms from age 7 to age 11.

No evidence was found in support of hypothesis 2.1. There was no statistically significant association between attainment grouping transition between ages 7 and 11 years and a less favourable change in parent- and teacher-reported externalising or internalising symptoms over the same period.

Objective 2.2: To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is related to an increase in socioeconomic inequalities in children's externalising and internalising psychological symptoms between ages 7 and 11 years.

No evidence was found in support of hypothesis 2.2. The three-way interaction between transition to attainment grouping, time (MCS5) and family income was not statistically significant at the 5% level. Therefore, it can be argued that analyses of this PhD provide evidence of no increase in socio-economic related differences in MCS participants' psychological symptoms during primary school due to transition to attainment grouping.

Objective 2.3: To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with an increased gap in externalising and internalising psychological symptoms between MCS participants with higher verbal skills and participants with lower verbal skills at primary school entry (age 5) in England and Wales.

Hypothesis 2.3 was not supported. The gap in externalising and internalising psychological symptoms for MCS participants with different levels of verbal skills at school entry (high versus low scores) did not widen among those who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups

6.9: Discussion

While there is some research on the association between attainment grouping and achievement test scores in the UK (Education Endowment Foundation, 2018a), little attention has been paid to the potential role of attainment grouping in primary school and children's externalising or internalising psychological symptoms. Similarly, little is known on whether this association varies by family income or baseline verbal skills.

6.9.1: Comparison to previous literature

Overall, the change in parent-reported or teacher- reported externalising or internalising symptoms was not associated with the transition to attainment grouping between ages 7 and 11 years. MCS participants experienced a similar decrease in both parent-reported and teacher-reported externalising symptoms from age 7 to age 11, independent of a transition to attainment grouping. A similar increase in both parent-reported and teacher-reported internalising symptoms was observed, regardless of whether MCS participants changed to attainment grouping or not.

This finding is difficult to compare with previous published findings in the UK and other developed countries for several reasons. First, previous studies in the UK focused on a range of academic-related non-cognitive outcomes in response to attainment grouping among secondary aged pupils in England. Such outcomes were academic self-concept and liking for school (Ireson et al., 2001, MacIntyre and Ireson, 2002, Ireson and Hallam, 2005, Ireson and Hallam, 2009). Only one unpublished study with a focus on children's psychological symptoms of primary school pupils in response to attainment grouping in primary school has been conducted (McDool, 2018). Second, previous published studies in the UK as well as in other developed countries have explored the association between each specific type of attainment grouping (streaming or literacy setting or numeracy setting) in isolation, whereas evidence on the practice of attainment grouping by UK schools shows that different attainment grouping practices in primary school may be practiced in tandem (see section 4.3.1.4).

Despite the difficulties in comparing the finding of no association between the transition to attainment grouping and the change in children's parent- and teacher-reported externalising or internalising symptoms , it can be argued that this chapter's main finding contradicts the statistically significant findings reported by McDool (2018). This study found that MCS participants who were taught in sets for mathematics at both age 7 and age 11 had higher teacher-reported internalising and parent-reported externalising psychological symptoms at the 10% significance level than MCS participants were not taught in sets for mathematics at both ages. This discrepancy could be attributed to differences in the definition of the exposure of interest; in this chapter's analyses attainment grouping reflects the combined effect of different types of between-class attainment grouping such as streaming or setting for literacy or mathematics, and this may have diluted the effect found by the study conducted by McDool. Despite the fact that McDool (2018) controlled for a wide set of time-invariant and time-varying confounders , the possibility of omitted time-varying confounding when using only fixed effects design could not be ruled out. In contrast, the fixed-effects difference-in-difference modelling employed by the analyses of this chapter could further minimise time-varying confounding bias. Fixed-effects difference-in-differences design is dependent on the assumption that confounders varying across the two groups (in this research those who changed to attainment grouping and those who did not) are time invariant, and time-varying confounders are group invariant, thus effectively minimising the risk of bias from spurious shocks over time. In addition, the different p-value thresholds interpreted as statistically significant (here significance level was set at the 5% level) should be considered.

The results of this chapter also indicated no significant differences or polarisation in children's parent- or teacher-reported externalising or internalising symptoms during primary school according to family income. A possible explanation might be that family income differences in externalising or internalising symptoms were already internalised by MCS participants from a very early age in life, which even dated before the start of primary school (Kelly et al., 2011), and simply were carried forward during primary school years, regardless of attainment grouping.

6.9.2: Strengths and limitations

This chapter's analyses contribute to the existing UK literature in several ways. Firstly, the effect of the transition to attainment grouping in primary school on externalising or internalising psychological symptoms was examined, using a large nationally representative sample of primary aged pupils in England and Wales. Secondly, it was attempted to overcome methodological issues in addressing unobserved confounding. The analyses of this chapter accounted for hard to measure fixed differences between children (i.e unobserved stable personality traits) by using each MCS participant as their own control. The difference-in-difference approach also accounted for all shared changes over time between MCS primary schools in England and Wales which chose to group their students by attainment between 7 and 11 years and those which did not. For instance, MCS primary schools in England and Wales which adopted attainment grouping between 7 and 11 years were exposed to the same external factors such as the Great recession and its aftermath (2008-2012).

This study also presents limitations. The lack of MCS teacher surveys for Scotland and Northern Ireland limited this chapter's analyses and therefore the extent to which the findings can be generalised beyond England and Wales. Non-response at MCS 5 teacher survey was assumed to be MCAR as it was the case for MCS 4 teacher survey. It might be possible MCS participants' teachers who responded or did not respond to the survey were different in terms of their individual or the school characteristics in which they were teaching at the time of MCS 5. For instance, if teachers in less well-resourced MCS primary schools in disadvantaged areas in England and Wales were less likely to respond to the MCS 5 survey, it would be possible that MCS participants with higher externalising or internalising psychological symptoms were not included for analyses of this chapter, therefore the association between the transition to attainment grouping and children's parent-reported or teacher-reported externalising or internalising psychological symptoms would be underestimated.

In summary, evidence was provided of association between the attainment grouping between 7 and 11 years and the change in MCS participants externalising or internalising psychological symptoms from 7 to 11 years. This association did not vary by the MCS participant's family income or baseline verbal skills. The next chapter will assess the association between attainment grouping at age 7and trajectories of parent-reported externalising and internalising symptoms from age 7 to age 14.

Chapter 7: Attainment grouping at age 7 and trajectories of externalising and internalising symptoms from age 7 to age 14

7.1: Chapter overview

The aim of this chapter is to explore the long-term association between attainment grouping at age 7 years and trajectories of externalising and internalising symptoms from childhood to adolescence (from 7 to 14 years of age). This chapter presents results of statistical analyses related to objectives 3.1 to 3.3.

Objective 3.1 aims to examine whether attainment grouping at age 7 is associated with a less favourable change in parent-reported externalising and internalising psychological symptoms. It is hypothesised that being exposed to attainment grouping at age 7 will be associated with a slower decrease in parent-reported externalising and a faster increase in parent-reported internalising psychological symptoms between ages 7 and 14 years, compared to mixed attainment grouping at age 7 which will be associated with a faster decrease in externalising psychological symptoms and a slower increase in internalising psychological symptoms between ages 7 and 14 years.

Objective 3.2 aims to test whether attainment grouping at age 7 is related to an increase in socioeconomic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years. It is hypothesised that attainment grouping at age 7 will be associated with an increase in socio-economic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years. Mixed grouping at age 7 will be associated with a smaller increase in socio-economic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.

Objective 3.3 aims to test whether attainment grouping at age 7 is associated with an increased gap in externalising and internalising psychological symptoms among MCS participants with higher and lower levels of verbal skills at age 5. It is hypothesised that the gap in externalising and internalising psychological symptoms will increase between age 7 and 14, contingent on differences in verbal skills at age 5 for those MCS participants in attainment groups at age 7. The gap in externalising and internalising psychological symptoms among MCS participants with higher and lower verbal skills at age 5 and who were taught in mixed attainment groups at age 7, will not increase as much as the gap for MCS participants with higher and lower verbal skills at age 5 and who were taught in attainment groups at age 7.

First, a detailed description of multiple imputation is given, and an account of confounding factors adjusted for in the multivariable growth curve models of this chapter is provided. Then a summary and a discussion of the findings in relation to research objectives 3.1 to 3.3 and their related hypotheses is given. The variables used in the empirical analyses of this chapter are presented in Table A4.1 in the Appendix 4.

7.2: Multiple imputation

7.2.1: Imputation model sample size

As previously described in chapter 4, during the MCS 4, teachers from primary schools across the UK provided information for 8,775 MCS participants. Multiple births (twins and triplets) were excluded (N=116) from further statistical analyses to remove family-level clustering. Information was then imputed for 8,659 MCS participants.

7.2.2: Item non-response

The percentage of missing values in either parent-reported internalising or externalising symptoms from age 7 to age 14 ranged from approximately 5% to approximately 22.5% as it can be seen in Table 7.1. The greatest amount of missing information was observed for teacher-reported school-related parameters. The highest percentage of missing information was observed for the number of years working as a teacher altogether. Approximately 44% of information on this variable was missing.

Using the recommended formula of *equation 4.1* which is described in detail in chapter 4. Fifty (50) multiple imputed datasets were generated, and the reproducibility seed was set at 2300.

Table 7. 1: Percentage of missing values in the outcomes, exposures, and confounders(N=8,659)

	Complete cases	**Missing (N)**	**Missing (%)**
Parent-reported SDQ externalising score (MCS 4	8,501	158	1.82%
Parent-reported SDQ externalising score (MCS 5)	7,454	1,205	13.92%
Parent-reported SDQ externalising score (MCS 6)	6,707	1,952	22.54%
Parent-reported SDQ internalising score (MCS 4)	8,498	161	1.86%
Parent-reported SDQ internalising score (MCS 5)	7,456	1,203	13.89%
Parent-reported SDQ internalising score (MCS 6)	6,710	1,949	22.51%
Attainment grouping (MCS 4)	8,026	633	7.31%
MCS strata	8,659	0	0.00%
School type (MCS 4)	7,713	946	10.93%
Number of classes in child's year (MCS 4)	5,015	3,644	42.08%
Years of teaching experience (MCS 4)	4,851	3,808	43.98%
Class size (MCS 4)	5,066	3,593	41.49%
Child in mixed-year group (MCS 4)	5,144	3515	40.59%
Number of children with EAL in child's classroom (MCS 4)	4,671	2,525	29.16%

	Complete cases	Missing (N)	Missing (%)
Number of children with SEN in child's classroom (MCS 4)	4,948	3,711	42.86%
Family income (MCS 4)	8,655	4	0.05%
Verbal skills at MCS 3	8,255	434	5.01%
Gender	8,659	0	0%

7.2.3: Imputation model specification

The specified imputation model of this chapter included information on attainment grouping at age 7, parent-reported externalising and internalising symptoms at ages 7, 11 and 14 years and time invariant information at age 7 as reported by the MCS participant's main carer and the child's teacher. Information was imputed on the type of school the child was attending , the number of classes in the child's school year, the number of pupils in child's classroom, the number of pupils with Special and Educational Needs (SEN) and English as an Additional Language (EAL)in child's classroom, whether the MCS participant was taught in a mixed-year group, whether there were disruptive pupils in child's ' class and information on the years of teaching working experience. A number of interactions were included in the imputation model; a) between attainment grouping at age 7 and child's age in years, and b) attainment grouping at age 7 and family income, and c) attainment grouping at age 7 and verbal skills at age 5.

Information on language spoken at home during each MCS interview and information on child's non-verbal reasoning ability at age 5 was included in the imputation model as auxiliary information to improve the prediction of imputed values on the parent-reported externalising and internalising symptoms from age 7 to age 14.Preliminary analysis undertaken found that these variables were strong predictors of response to parent reported SDQ at MCS 4, MCS 5 and MCS 6.

Information on the MCS stratum was included to account for the stratified survey design of MCS, while survey design and attrition weights were included. An interaction term between survey design, attrition weights and information on MCS stratum was also included in the imputation model in order to take into account the unequal probabilities of selection in the MCS, the loss-to follow up, and ascertain that imputed values were not unduly influenced by subjects from groups that were oversampled in the MCS (Seaman et al., 2012).

7.2.4: Imputation diagnostics

Statistical checks using Stata's *midiagplots* (Eddings and Marchenko, 2012) on the 50 multiple imputed datasets revealed that multiple imputation produced reasonable values of imputed data. (The imputation models did respect the bounds, all the imputed values for both parent-reported externalising or internalising scores lay between 0 and 20).

The average within-imputation and between-imputation variances for all the other variables included in the imputation models were small and the relative efficiency of the imputed values exceeded 99%. There was an acceptable amount of Monte Carlo Error (MCE): MCE for all coefficients was less than 10% of their standard error, the MCE of the coefficients' T-statistic was less than 0.1 and finally MCE of all coefficients p-value was less than 0.01 if p-value for a coefficient was $p<0.05$ and below 0.02 if p-value for coefficients with $p<0.10$ as it has been suggested (White et al., 2011). MCE was unlikely to have affected the conclusions reached in the analysis stage. In the analysis stage estimates from each imputed dataset were combined under Rubin's rules (Rubin, 1987) by using Stata's command *mi estimate*.

7.3: Analysis sample

Teachers in UK primary schools returned questionnaires for 8,775 MCS participants at MCS4. Twins and triplets (N=116) were excluded from statistical analysis due to their small number and the possibility that their social and emotional development is determined differently (Sutcliffe and Derom, 2006). Thirty (N=30) MCS participants with missing information on all time points (MCS4, MCS5 and MCS6) were excluded from further statistical analysis. Descriptive, bi-variable, and multivariable growth curve analysis was conducted in multiple imputed data of 8,629 MCS participants. Analysis was also conducted in 3,853 MCS participants with complete data on parent-reported externalising/internalising symptoms, attainment grouping at age 7 and time-invariant school-related confounders at age 7. This was done to explore whether the results using multiple imputed data were different from the ones obtained from complete cases.

7.4: Analytical model specification

Provided that three repeated measures (MCS 4, MCS 5 and MCS 6) over-identify a linear trajectory (that is, there is more observed information than estimated information) (Curran et al., 2010) a linear growth model was considered with cross-level interactions[13] where

$$\boldsymbol{y_{ij} = \beta_0 + \beta_{1age_{ij}} + e_{ij}}$$

$$\boldsymbol{\beta_{0j} = \beta_0 + \beta_2 X_j + u_{0j}}$$

$$\boldsymbol{\beta_{1j} = \beta_1 + \beta_3 X_j + u_{1j}}$$

$$\boldsymbol{\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right\}, e_{ij} \sim N\left(0, \sigma^2_e\right)}$$

The mixed-effects formulation of the model with cross-level interactions can also be written as;

$$\boldsymbol{y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 X_j + \beta_3 X_j \times age_{ij} + u_{0j} + u_{1jage_{ij}} + e_{ij}}$$

[13] An interaction between time (measured as the average age in MCS 4 to MCS 6) and attainment grouping at age 7 can be referred as a cross-level interaction since attainment grouping represents a pedagogic practice decided at the school-level (higher level), while time refers to the changes in externalising and internalising difficulties at the individual level (lower level) across MCS4 to MCS6.

Equation 7. 1

Where $y_{ij} = \beta_0 + \beta_{1age_{ij}} + \beta_2 X_j + \beta_3 X_j \times age_{ij}$ is the fixed part of the model and $u_{0j} + u_{1jage_{ij}} + e_{ij}$ is the random part of the model. A detailed description of the model is given below with.

$\boldsymbol{y_{ij}}$= parent-reported externalising or internalising SDQ score for the ith individual in the jth group

$\boldsymbol{\beta_0}$**=** overall intercept

$\boldsymbol{\beta_{1age_{ij}}}$ =overall slope for linear time

$\boldsymbol{\beta_2 X_j}$ =explanatory variable for which values vary between individuals within group

$\boldsymbol{\beta_3 X_j \times age_{ij}}$**=**Cross-level interaction between an explanatory variable and linear function of time.

The cross-level interaction will explain away the random-slope variation, while the explanatory variable main effect will explain away the random-intercept variation.

$\boldsymbol{\beta_{0j}}$ =intercept (mean of *y*) for a given group *j*

$\boldsymbol{e_{ij}}$ =-individual level residuals (difference between y_{ij} and β_0)

$\boldsymbol{u_{0j}}$=group level residuals (difference between β_{0j} and β_0)

$\boldsymbol{u_{1jage_{ij}}}$ = group level residuals (difference between β_{0j} and β_0) across age

$\boldsymbol{\sigma^2_{u0}}$= group level between variance

$\boldsymbol{\sigma^2_e}$=individual level (within group) variance.

Linear growth curve models were then fitted in STATA v.15 using the *xtmixed* command. All repeated observations at MCS 4, MCS 5 and MCS 6 were clustered within cohort members (*mcsid:*). Time was parameterized using children's average age in years at survey time and centred at age 7 (baseline) which allowed the intercept to be interpreted as the mean value of parent-reported mean externalising and internalising symptoms at the baseline (time=0), with subsequent measures at times 4 and 7 representing parent-reported externalising and internalising symptoms at ages 11 and 14 years respectively. Further, it was assumed that not all individuals have the same slope for age, therefore the coefficient for age could vary across individuals. In doing so, the variance in mean externalising or internalising symptoms and within-individual correlation could depend on children's average age in years. Therefore, the *xtmixed* command was modified to include a random slope for time after *mcsid: (mcsid:age)*.The *cov (unstruc)* option was also used since it specifies the form of the covariance matrix for the intercept and slope random effects. The default is for intercepts and slopes to be independent which is usually an unreasonable assumption to be made. A bivariate normal distribution for $u_{0j}$and u_{1jage_j} implies an ''unstructured'' covariance matrix.

7.4.1: Statistical approach

Bi-variable analysis was conducted in multiple imputed data to assess unadjusted relationships between the probability of the MCS participant being grouped by attainment at age 7 and parent-reported or teacher-reported school-related characteristics. Further, potential unadjusted associations between parent-reported or teacher-reported school-related characteristics and mean parent-reported externalising and internalising symptoms at age 7 were assessed through linear regression modelling. For ease of interpretation, candidate continuous confounding factors were categorised into tertiles for all the bi-variable analyses.

Potential associations between early attainment grouping and the development of mean parent-reported externalising and internalising symptoms from age 7 to age 14 as well as differential effects of early attainment grouping by family income and initial verbal skills were explored through growth curve modelling.

7.5: Imputed descriptive results

7.5.1: Parent reports of children’s externalising and internalising symptoms at ages 7, 11 and 14 years.

The distributions of parent-reported symptoms of externalising (Figure7.1) and internalising symptoms (Figure 7.2) at ages 7, 11 and 14 years were skewed to the right (positively skewed) with higher percentages of lower scores at each age. It should be noted that although parent-reported externalising and internalising symptoms were not normally distributed, there is evidence that standard linear modelling is useful in ‘’sufficiently large’’ samples (N>500) to analyse differences in trends even with non-normally distributed data (Lumley et al., 2002).

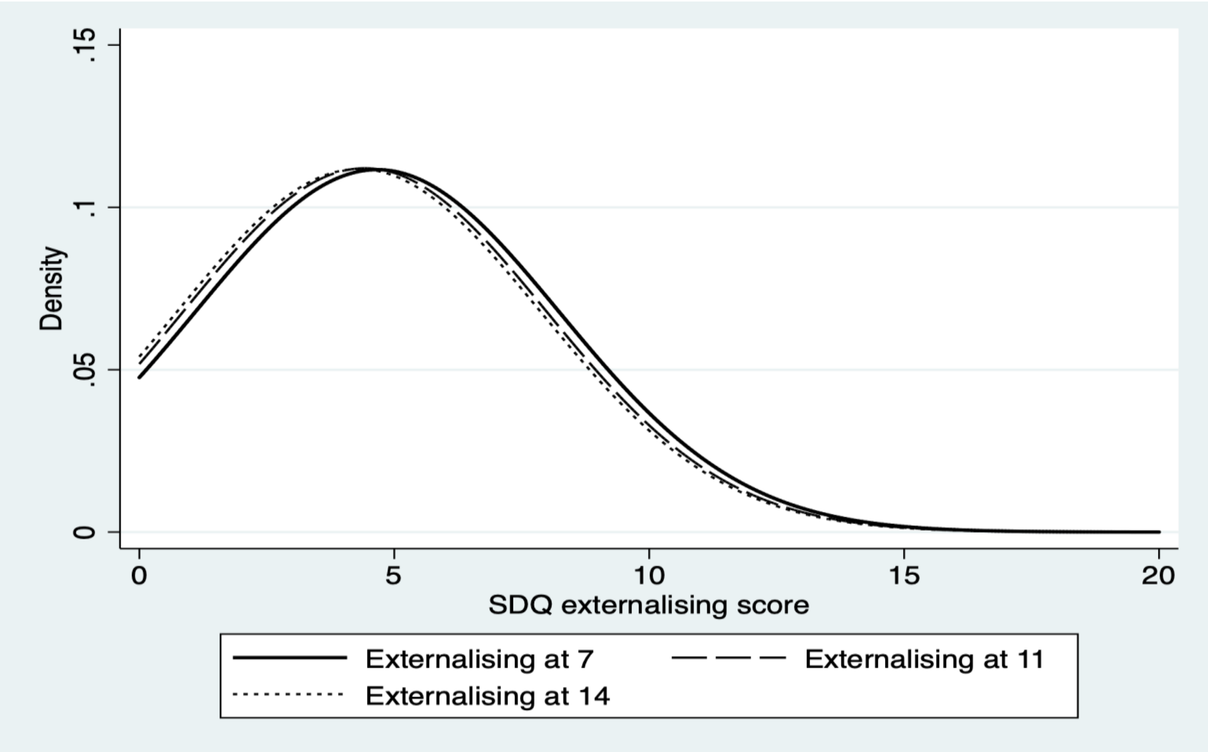

Figure 7. 1: Distribution of parent-reported externalising symptoms at ages 7, 11 and 14 years (N=8,629)

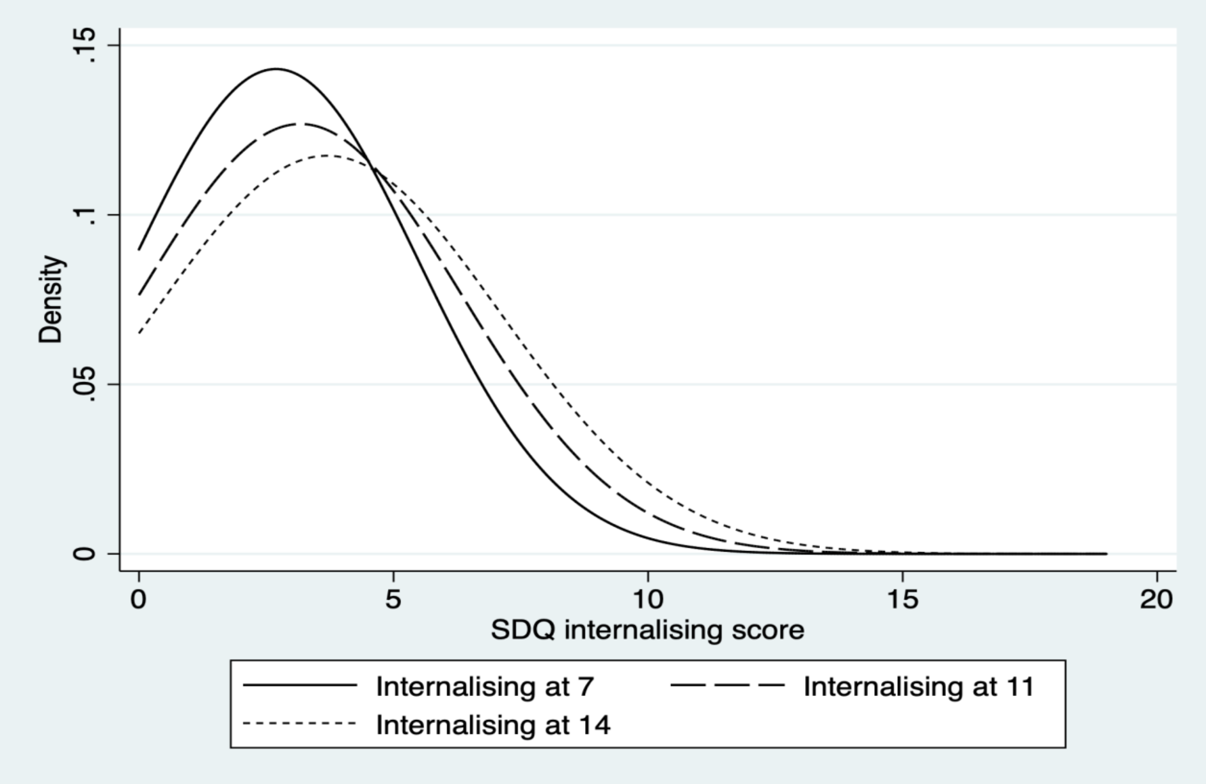

Figure 7. 2: Distribution parent-reported internalising symptomsat age 7, 11 and 14 (N=8629)

Figure 7.3 shows that children's mean externalising symptoms as reported by the child's parent decreased slightly from age 7 to age 14. Figure 7.4 highlights that children's mean symptoms of internalising difficulties increased substantially from a mean of 2.58 at age 7 to a mean 3.58 at age 14 (see Figure 7.4).

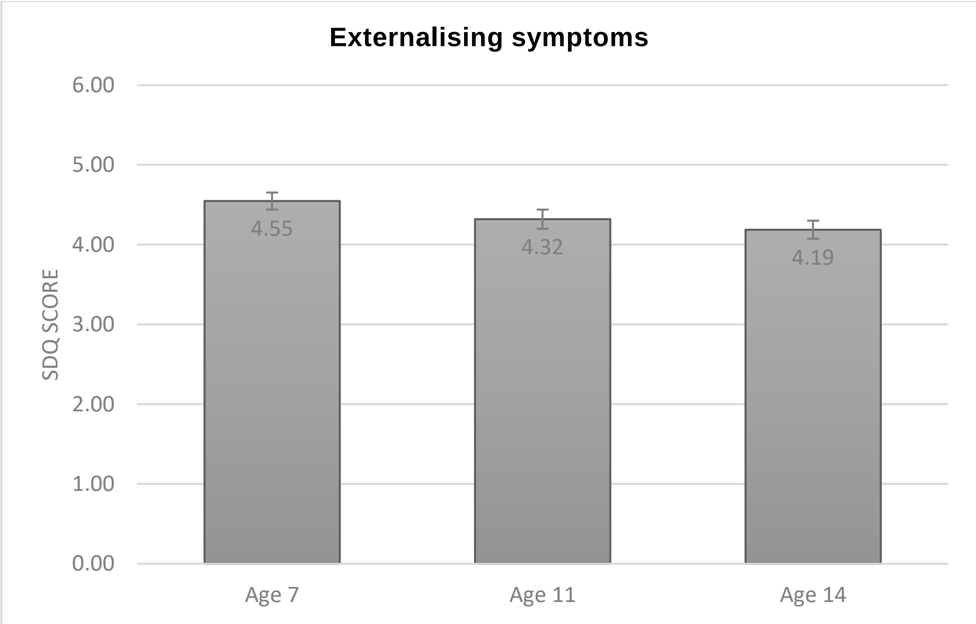

Figure 7. 3: Mean parent-reported externalising symptoms at ages 7, 11 and 14 years (N=8,629)

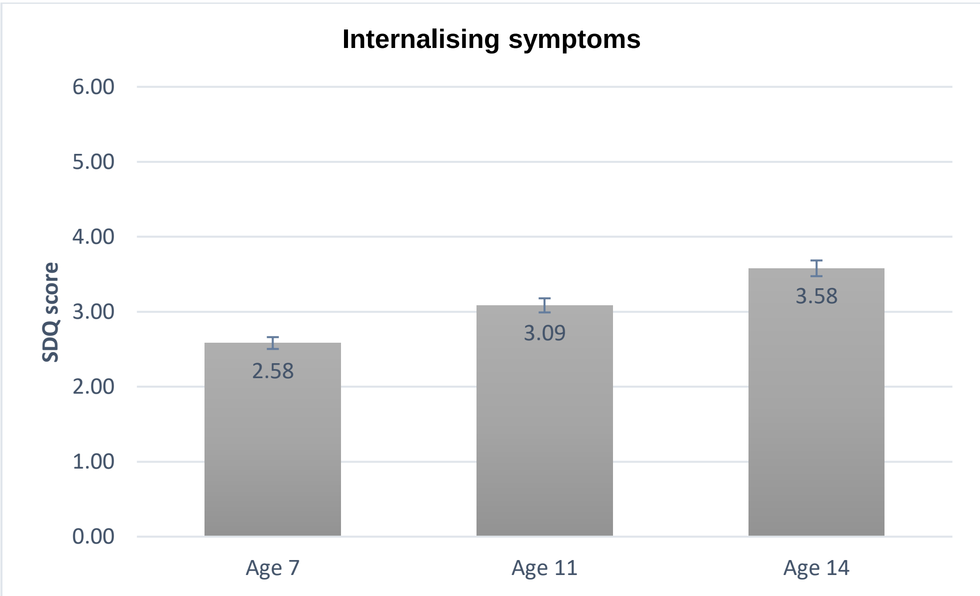

Figure 7. 4: Mean parent-reported internalising symptoms at ages 7, 11 and 14 years (N=8,629)

7.5.2: Socio-demographic characteristics of the analytic sample

Table 7.2 shows that the analytic sample was comprised of approximately an equal proportion of boys and girls and a very high proportion of MCS participants who were living in English advantaged areas (55.7%).

Table 7. 1: Basic socio-demographic characteristics of the analytic sample at age 7 (N=8,629)

	Percentage	S.E	95% CI	95% CI
Categorical variables				
CM's gender				
Male	49.8%	0.01	48.6%	51.1%
Female	50.2%	0.01	48.9%	51.4%
MCS strata				
England-Advantaged	55.7%	0.02	52.3%	59.1%
England - Disadvantaged	24.3%	0.01	21.5%	27.0%
England - Ethnic	3.2%	0.00	2.3%	4.0%
Wales - Advantaged	2.6%	0.00	2.0%	3.3%
Wales - Disadvantaged	2.0%	0.00	1.6%	2.4%
Scotland - Advantaged	5.7%	0.00	4.8%	6.6%
Scotland - Disadvantaged	3.0%	0.00	2.5%	3.6%
Northern Ireland - Advantaged	2.1%	0.00	1.6%	2.5%
Northern Ireland - Disadvantaged	1.4%	0.00	1.2%	1.6%

7.5.3: School-related characteristics of the analytic sample

Table 7.3 shows that 41.6% of children in the analytic sample were grouped by attainment at age 7, 68.2% were attending state-maintained primary schools and they were attending primary schools which on average had 2 classes in child's school year. Further, children in the analytic sample were taught on average in classrooms of 25 pupils when they were 7 years-old, they were also taught on average alongside one child with SEN and two children with EAL, whereas 33% of teachers responded that there were children in CMs classroom who displayed disruptive behaviour. Finally, teachers in the analytic sample were teaching altogether for an average of 14 years.

Table 7. 2: School-related characteristics at age 7 (N=8,629)

	Percentage	**S.E**	**95% CI**	**95% CI**
Categorical variables				
Attainment grouping at age 7				
Not grouped by attainment	58.4%	0.01	56.0%	60.9%
Grouped by attainment	41.6%	0.01	39.1%	44.0%
Type of school				
State-maintained	68.2%	0.01	65.6%	70.8%
Fee-paying	5.1%	0.01	3.7%	6.5%
Faith school	26.7%	0.01	24.4%	29.0%
Child in mixed-year class				
No	71.3%	0.01	69.4%	73.2%
Yes	28.7%	0.01	26.8%	30.6%
Disruptive peers in child's classroom				
No	66.7%	0.01	65.1%	68.3%
Yes	33.3%	0.01	31.7%	34.9%
Numerical variables				
	Mean	**S.E**	**95% CI**	**95% CI**
Number of classes in child's year	1.88	0.02	1.85	1.92
Number of children in child's class	25.41	0.10	25.2	25.61
Number of SEN children in child's class	1.40	0.04	1.33	1.48
Number of EAL children in child's class	2.01	0.12	1.78	2.25
Average years of teaching experience	13.62	0.17	13.28	13.96

7.6: Bi-variable analysis results

Figure 7.5 shows that MCS participants who were grouped by attainment at age 7 displayed slightly higher parent-reported externalising symptoms from age 7 to age 14. In addition, MCS participants who were grouped by attainment at age 7 had slightly higher internalising symptoms from age 7 to age 14 compared to MCS participants who were taught in mixed attainment classes.

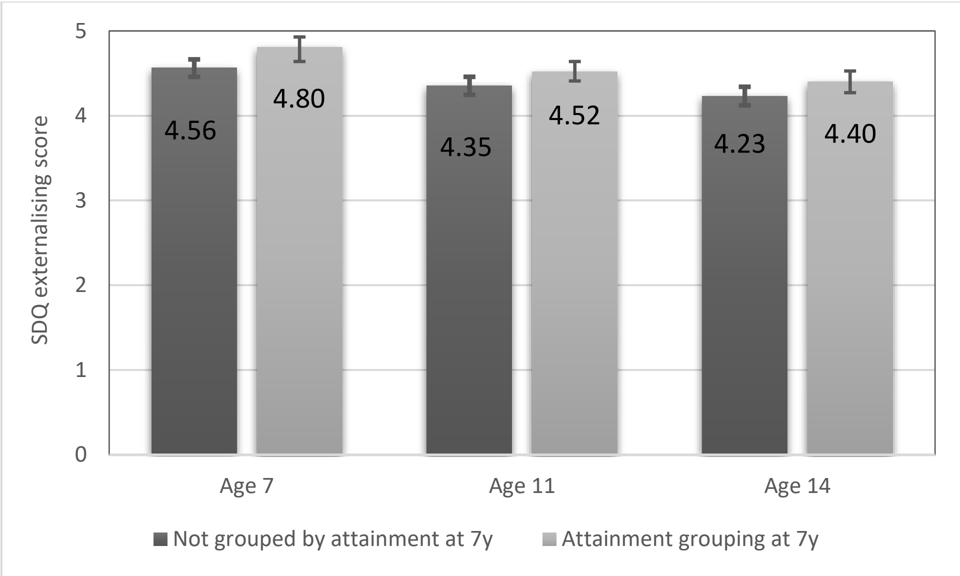

Figure 7. 5: Mean parent-reported externalising symptoms at ages 7,11 and 14, by attainment grouping status at age 7 (N=8,629)

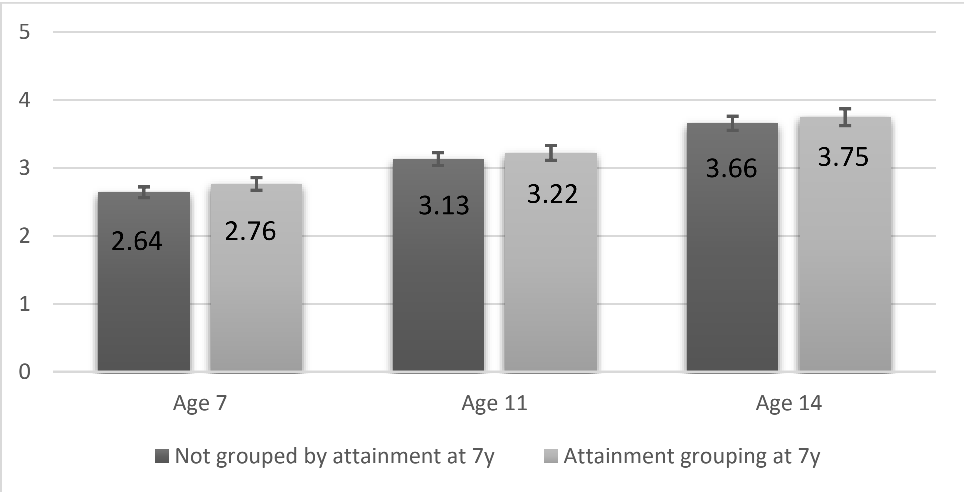

Figure 7. 6: *Mean parent-reported externalising symptoms at ages 7,11 and 14, by attainment grouping status at age 7 (N=8,629)*

Table 7.4 displays a lower probability of being grouped by attainment at age 7 if MCS participants were attending fee-paying and faith primary schools compared to children who were attending state-maintained primary schools in the UK. There was also a higher probability of being grouped by attainment at age 7 if children were attending larger primary schools, if they were taught in classrooms with a higher number of pupils, if there was a higher number of pupils with SEN or EAL, and if MCS participants were also taught in classrooms which included pupils with disruptive behaviour. A lower likelihood of being grouped by attainment at age 7 was associated with having more experienced teachers. No association was found between mixed-year group teaching and the probability of being grouped by attainment at age 7.

Table 7. 3: Percentages and odds ratios of being grouped by attainment at age 7 by school-related characteristics (N=8,629)

	%	**OR**	**S.E**	**95% CI**	**95% CI**
Variables					
Type of school					
State-maintained (Ref.)	44.2%				
Fee-paying	33.2%	0.63*	0.11	0.44	0.89
Faith school	38.0%	0.77*	0.08	0.62	0.96
No disruptive peers in child's classroom (Ref.)	40.2%				
Disruptive peers in child's classroom	45.8%	1.26**	0.10	1.08	1.46
Number of classes in child's year					
Lowest tertile (Mean=1) (Ref.)	31.7%				
Medium tertile (Mean=2)	46.2%	1.85***	0.14	1.60	2.15
Highest tertile (Mean=3)	52.3%	2.36***	0.23	1.95	2.85
Number of children in child's class					
Lowest tertile (Mean=20) (Ref.)	39.6%				
Medium tertile (Mean=27)	42.5%	1.13	0.09	0.97	1.31
Highest tertile (Mean=30)	44.5%	1.22*	0.10	1.04	1.44
Child in single-year group (Ref.)	42.4%				
Child in mixed-year group	41.4%	0.96	0.08	0.82	1.12
Number of SEN children in child's class					
Lowest tertile (Mean=0) (Ref.)	38.7%				
Medium tertile (Mean=1)	44.0%	1.25**	0.09	1.07	1.45
Highest tertile (Mean=5)	47.9%	1.45***	0.13	1.23	1.72
Number of EAL children in child's class					

	%	OR	S.E	95% CI	95% CI
Lowest tertile (Mean=0) (Ref.)	40.2%				
Medium tertile (Mean=1)	42.8%	1.12	0.17	0.95	1.31
Highest tertile (Mean=8)	45.4%	1.24*	0.01	1.05	1.46
Years of teaching experience					
Lowest tertile (Mean=3.8 years) (Ref.)	45.0%				
Medium tertile (Mean=12.6 years)	41.7%	0.88	0.07	0.75	1.02
Highest tertile (Mean=22.7 years)	39.0%	0.78**	0.06	0.67	0.91

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** $p<0.001$, ** <0.01, * <0.05

Tables 7.5 and 7.6 showed statistically significant associations between higher reported externalising and internalising symptoms at age 7 and attendance to state-maintained UK primary schools, higher numbers of pupils with SEN and EAL in children's classrooms and for children who had classmates with disruptive behaviour. In contrast, there was a statistically significant association between lower parent-reported externalising and internalising symptoms at age 7, and greater class size. Considering the results of the logistic and linear bi-variable analyses, multivariable growth curve models were adjusted for school type, the number of pupils in child's classroom, the number of pupils with SEN and EAL in child's classroom and whether there were pupils with disruptive behaviour in child's classroom.

Table 7. 4: The association between mean parent-reported externalising symptoms at age 7 and school-related characteristics (N=8,629)

	Mean	B	S.E	95% CI	95% CI
Variables					
Type of school					
State-maintained (Ref.)	4.98				
Fee-paying	3.66	-1.32***	0.18	-1.68	-0.97
Faith school	4.46	-0.52***	0.11	-0.74	-0.30
No disruptive peers in child's classroom (Ref.)	4.55				
Disruptive peers in child's classroom	5.27	0.72***	0.12	0.48	0.96
Number of classes in child's year					
Lowest tertile (Mean=1) (Ref.)	4.79				
Medium tertile (Mean=2)	4.79	0.00	0.13	-0.25	0.25
Highest tertile (Mean=3)	4.82	0.04	0.15	-0.26	0.33
Number of children in child's class					
Lowest tertile (Mean=20) (Ref.)	5.00				
Medium tertile (Mean=27)	4.84	-0.16	0.15	-0.45	0.13
Highest tertile (Mean=30)	4.56	-0.44***	0.15	-0.73	-0.14
Child in single-year group (Ref.)	4.70				
Child in mixed-year group	5.04	0.34**	0.13	0.09	0.58
Number of SEN children in child's class					
Lowest tertile (Mean=0) (Ref.)	4.58				
Medium tertile (Mean=1)	4.82	0.23	0.13	-0.02	0.49
Highest tertile (Mean=5)	5.30	0.72***	0.17	0.39	1.05
Number of EAL children in child's class					
Lowest tertile (Mean=0) (Ref.)	4.67				

	Mean	B	S.E	95% CI	95% CI
Medium tertile (Mean=1)	4.76	0.08	0.15	-0.22	0.39
Highest tertile (Mean=8)	5.04	0.36**	0.12	0.12	0.61
Years of teaching experience					
Lowest tertile (Mean=3.8 years) (Ref.)	4.83				
Medium tertile (Mean=12.6 years)	4.77	-0.05	0.13	-0.31	0.20
Highest tertile (Mean=22.7 years)	4.78	-0.05	0.13	-0.31	0.20

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table 7. 5: The association between mean parent-reported internalising symptoms at age 7 and school-related characteristics (N=8,629)

Variables	**Mean**	**B**	**S.E**	**95% CI**	**95% CI**
Type of school					
State-maintained (Ref.)	2.88				
Fee-paying	1.77	-1.11***	0.15	-1.41	-0.81
Faith school	2.54	-0.34***	0.09	-0.51	-0.16
No disruptive peers in child's classroom (Ref.)	2.59				
Disruptive peers in child's classroom	3.03	0.43***	0.10	0.23	0.64
Number of classes in child's year					
Lowest tertile (Mean=1) (Ref.)	2.77				
Medium tertile (Mean=2)	2.75	-0.02	0.10	-0.23	0.19
Highest tertile (Mean=3)	2.68	-0.09	0.12	-0.34	0.16
Number of children in child's class					
Lowest tertile (Mean=20) (Ref.)	2.91				
Medium tertile (Mean=27)	2.72	-0.19	0.11	-0.41	0.03
Highest tertile (Mean=30)	2.59	-0.32**	0.11	-0.53	-0.11
Child in single-year group (Ref.)	2.73				
Child in mixed-year group	2.79	0.06	0.10	-0.14	0.27
Number of SEN children in child's class					
Lowest tertile (Mean=0) (Ref.)	2.66				
Medium tertile (Mean=1)	2.72	0.06	0.09	-0.12	0.23
Highest tertile (Mean=5)	3.00	0.34**	0.12	0.10	0.58
Number of EAL children in child's class					
Lowest tertile (Mean=0) (Ref.)	2.63				

Variables	Mean	B	S.E	95% CI	95% CI
Medium tertile (Mean=1)	2.58	-0.05	0.12	-0.29	0.19
Highest tertile (Mean=8)	3.04	0.41***	0.11	0.20	0.62
Years of teaching experience					
Lowest tertile (Mean=3.8 years) (Ref.)	2.85				
Medium tertile (Mean=12.6 years)	2.69	-0.16	0.10	-0.36	0.05
Highest tertile (Mean=22.7 years)	2.67	-0.18	0.11	-0.40	0.04

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

A gradient was apparent in children's parent-reported externalising symptoms at age 7 by family income (see Figure 7.7). This illustrates that for each decrease in the family income quantile, starting from the highest income quantile, there was a statistically significant increase in parent-reported externalising symptoms at age 7.

Although children from low-income families had higher parent-reported internalising symptoms at age 7 compared to children from high-income families, there was a less clear gradient of an increase by family income. There was an overlap of the 95% CIs for quantiles 2 & 3 and the lowest quantiles for internalising symptoms (see Figure 7.8).

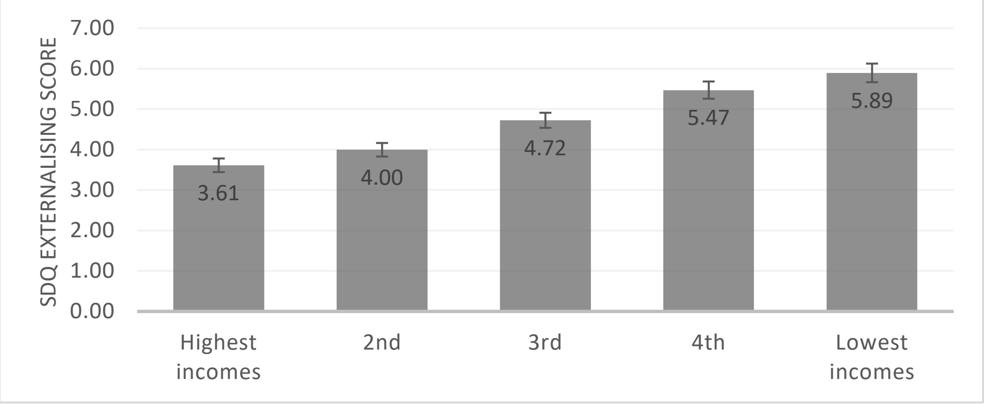

Figure 7. 7 : Mean parent-reported externalising symptoms by family income at MCS 4

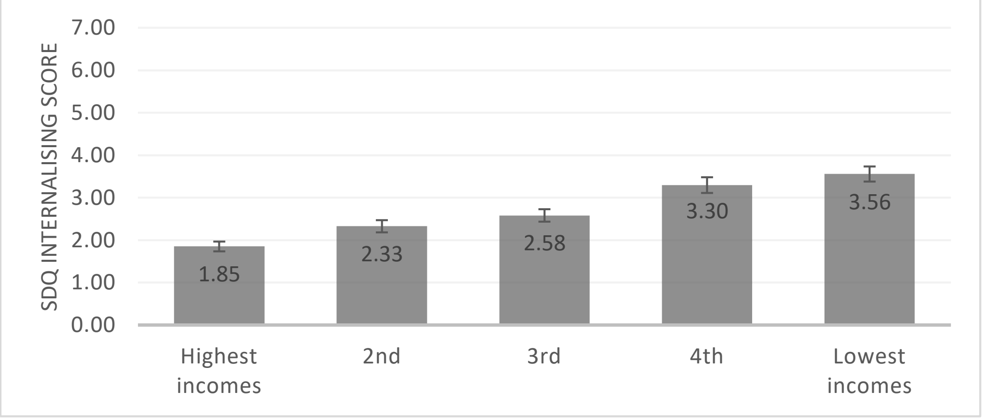

Figure 7. 8: parent-reported internalising symptoms by family income at MCS 4

7.7: Growth Curve Modelling results

7.7.1: Attainment grouping at age 7 and the development of externalising symptoms from age 7 to age 14

Table 7.8 highlights a small, but statistically significant positive association between attainment grouping at age 7 and parent-reported externalising symptoms at age 7 (β=0.19) in the minimally adjusted growth curve model (Model 1) using multiple imputed data. The interaction term between attainment grouping at age 7 and MCS participant's age showed moderate stability from age 7 to age 14, hence the small and non-significant interaction coefficient. Model 1 showed that MCS participants who were living in English disadvantaged and ethnically diverse areas as well as in Welsh, Scottish and Northern Irish areas had higher parent-reported externalising symptoms at age 7 compared to MCS participants who were residing in English advantaged areas. In contrast, MCS participants who were living in advantaged areas in Northern Ireland had lower parent-reported externalising symptoms than their peers in advantaged areas in England.

The random effects coefficients of model 1 also showed that there was a considerable variability in parent-reported reported symptoms of externalising symptoms at age 7 which increased with the passage of time.

Model 2, which represents the fully adjusted growth curve model, suggested an attenuation of the relationship between early attainment grouping and parent-reported externalising symptoms at baseline/intercept after adjustment for school-related confounders, rendering the association of interest not statistically significant at the $p<0.05$ level. Evidence was then provided of no association between attainment grouping at age 7 and trajectories of psychological symptoms from childhood to adolescence (see also Figure 7.9). It was also observed that adjustment for school-related confounders attenuated the observed associations between ethnic and disadvantaged MCS strata in England, Wales and Scotland compared to the advantaged

MCS stratum in England and parent-reported externalising symptoms. This means that non-adjustment for school-related confounding could inflate differences between children living in advantaged areas in England and those children living in disadvantaged and ethnically diverse areas in England. Therefore, non-adjustment for school-related confounding could falsely lead to the conclusion that there is a statistically significant association between living in disadvantaged areas in Wales, Scotland and Northern Ireland, and higher mean parent-reported externalising symptoms. On the contrary, failure to consider school-related confounders might underestimate differences between children living in advantaged areas in Scotland and Northern Ireland compared to children living in advantaged areas in England regarding their reported externalising symptoms.

Statistically significant associations were found between attending fee-paying and faith primary schools in the UK and lower mean reported externalising symptoms. A bigger classroom size was also associated with lower parent-reported externalising symptoms. Besides, higher parent-reported externalising symptoms were associated with having peers with disruptive behaviour in the same classroom and being in a classroom with a greater number of classmates with SEN. Turning now to bias and efficiency of growth curve modelling estimates, Tables 7.7 and 7.8 show the results of the main analysis using multiple imputation and complete cases. For all school-related confounders, including attainment grouping at age 7, the coefficients were similar, whereas the 95% CI were narrower and standard errors were smaller for analyses using for multiple imputed data, demonstrating the efficiency gain and less biased results of multiple imputation over complete cases.

Table 7. 6: Fixed and random effects of the partially (Model 1) and the fully adjusted (Model 2) growth curve modelling of parent-reported externalising symptoms (multiple imputed data)

	Model 1-Multiple imputation (n=8,629)				**Model 2-Multiple imputation (n=8,629)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.19*	0.08	0.03	0.35	0.15	0.07	-0.01	0.31
Age	-0.17***	0.03	-0.22	-0.12	-0.18***	0.02	-0.22	-0.14
Attainment grouping at age 7* Age	-0.04	0.04	-0.12	0.04	-0.04	0.04	-0.12	0.04
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.91***	0.09	0.72	1.09	0.75***	0.10	0.56	0.94
England - Ethnic	0.71***	0.13	0.45	0.96	0.53**	0.17	0.20	0.86
Wales - Advantaged	0.00	0.17	-0.34	0.34	-0.28	0.18	-0.63	0.07
Wales - Disadvantaged	0.68***	0.13	0.43	0.94	0.22	0.14	-0.06	0.50
Scotland - Advantaged	-0.24	0.15	-0.53	0.05	-0.51**	0.15	-0.80	-0.22
Scotland - Disadvantaged	0.39*	0.16	0.08	0.70	0.15	0.16	-0.16	0.46
Northern Ireland - Advantaged	-0.34*	0.17	-0.68	-0.01	-0.66***	0.18	-1.02	-0.30
Northern Ireland - Disadvantaged	0.45**	0.15	0.15	0.75	0.23	0.18	-0.12	0.57
Female	-1.13***	0.04	-1.26	-1.00	-1.12	0.07	-1.25	-0.98
School type (Ref. State-maintained)								
Fee-paying					-1.41	0.20	-1.80	-1.02

	Model 1-Multiple imputation (n=8,629)				Model 2-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Faith					-0.37	0.09	-0.54	-0.19
Classroom size					-0.05	0.01	-0.07	-0.03
Disruptive peers in CM's class					0.38	0.010	0.19	0.58
Number of pupils with SEN in CM's class					0.08	0.02	0.04	0.11
Number of pupils with EAL in CM's class					0.00	0.01	-0.01	0.02
Constant	4.20***	0.07	4.06	4.34	5.59	0.24	5.11	6.06
Random-effects Parameters								
Between-child intercept variance	3.10***	0.03	3.04	3.16	3.06***	0.03	3.00	3.12
Between-child slope variance	0.96***	0.03	0.91	1.02	0.96***	0.03	0.91	1.02
Between-child intercept and slope covariance	-0.30***	0.02	-0.34	-0.26	-0.30***	0.02	-0.34	-0.26
Between-occasion variance	1.79***	0.02	1.76	1.82	1.79***	0.02	1.76	1.82

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05, ^{NS} Random effect not significant*

Table 7. 7: Fixed and random effects of the partially (Model 1) and the fully adjusted (Model 2) growth curve modelling of parent-reported externalising symptoms (complete cases)

	Model 1-Complete case analysis (n=3,853)				**Model 2-Complete case analysis (n=3,853)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.21	0.12	-0.02	0.44	0.17	0.12	-0.04	0.38
Age	-0.19***	0.04	-0.26	-0.12	-0.19***	0.04	-0.24	-0.13
Attainment grouping at age 7*age	0.00	0.06	-0.11	0.12	0.01	0.06	-0.10	0.11
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.92***	0.13	0.66	1.18	0.74***	0.14	0.47	1.00
England - Ethnic	0.51**	0.19	0.15	0.87	0.36	0.21	-0.06	0.78
Wales - Advantaged	0.38	0.28	-0.17	0.92	0.09	0.28	-0.46	0.63
Wales - Disadvantaged	0.45*	0.19	0.07	0.82	-0.04	0.20	-0.44	0.35
Scotland - Advantaged	0.04	0.22	-0.38	0.46	-0.30	0.22	-0.73	0.12
Scotland - Disadvantaged	0.46*	0.22	0.04	0.89	0.14	0.22	-0.29	0.56
Northern Ireland - Advantaged	-2.67	3.20	-8.94	3.60	-2.98	3.16	-9.18	3.21
Northern Ireland - Disadvantaged	1.92	1.32	-0.69	4.52	1.66	1.31	-0.91	4.23
Female	-1.27***	0.10	-1.47	-1.07	-1.24***	0.10	-1.43	-1.04
School type (Ref. State-maintained)								
Fee-paying					-1.63***	0.24	-2.10	-1.16
Faith					-0.48***	0.12	-0.72	-0.24
Classroom size								
Disruptive peers in child's class					0.35**	0.11	0.13	0.56
Number of pupils with SEN in child's class					0.07**	0.02	0.02	0.11

	Model 1-Complete case analysis (n=3,853)				Model 2-Complete case analysis (n=3,853)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Number of pupils with EAL in child's class					0.00	0.01	-0.02	0.02
Constant	4.22***	0.10	4.01	4.42	5.83***	0.30	5.24	6.41
Random-effects Parameters								
Between-child intercept variance	0.96***	0.04	0.88	1.05	3.08***	0.05	2.99	3.17
Between-child slope variance	3.12***	0.05	3.03	3.21	0.96***	0.04	0.88	1.05
Between-child intercept and slope covariance	-0.25***	0.03	-0.31	-0.19	-0.25***	0.03	-0.31	-0.19
Between-occasion variance	1.80***	0.02	1.76	1.85	1.804***	0.02	1.758	1.850

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05, ^{NS} Random effect not significant*

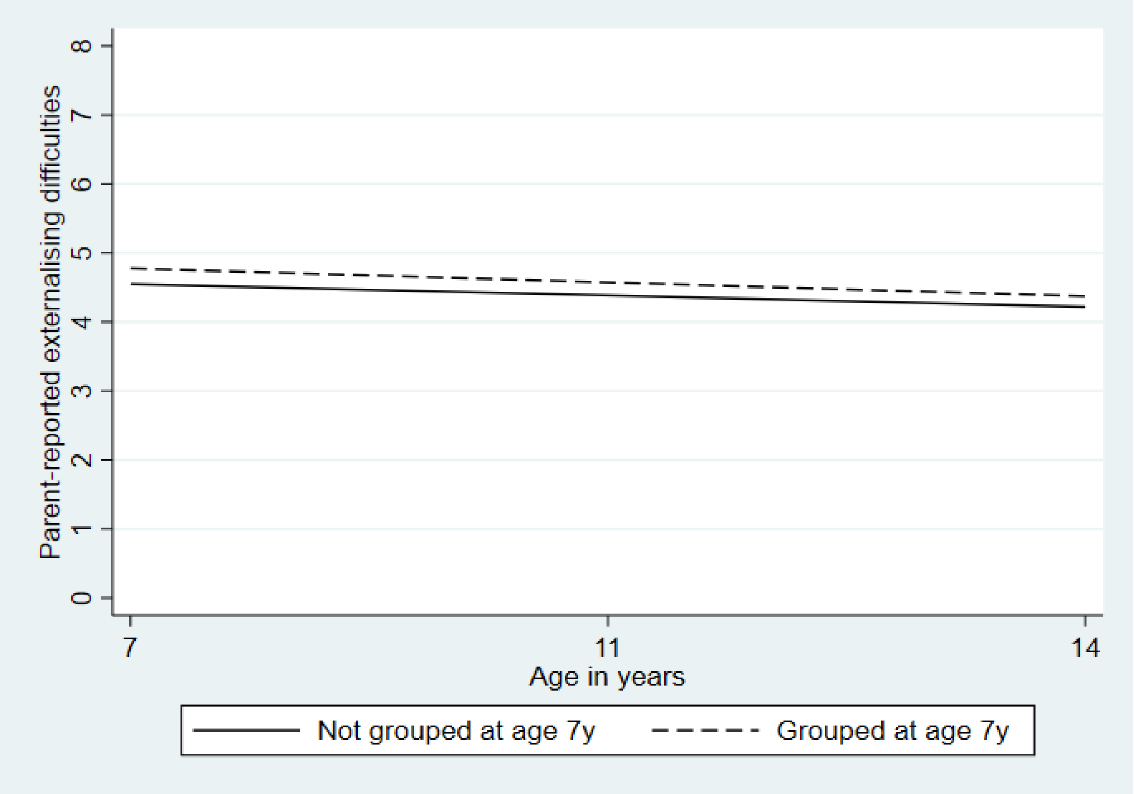

Figure 7. 9: Attainment grouping at age 7y and trajectories of parent-reported externalising symptoms from 7 to 11 years of age (N=8,629).

7.7.2: Attainment grouping at age 7 and trajectories of externalising symptoms from age 7 to age 14 according to family income

Model 4 (Table 7.9 and Figure 7.10), which tested interactions between attainment grouping at age 7, family income and age did not provide evidence of effect modification by family income at age 7 and from age 7 to age 14 in either complete cases or multiple imputed data. This suggests that family income-related differences in the development of externalising symptoms between ages 7 and 14 years were unaffected by exposure to early attainment grouping.

Table 7. 8: Fixed and random effects for Model 4; the association between attainment grouping at age 7, family income and parent-reported externalising symptoms

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	-0.07	0.21	-0.48	0.35	-0.01	0.15	-0.31	0.29
Age	-0.18***	0.03	-0.24	-0.13	-0.18***	0.02	-0.22	-0.14
Attainment grouping at age 7 * age	-0.10	0.11	-0.33	0.13	-0.07	0.08	-0.23	0.09
Family income (Ref. Highest incomes								
2nd quantile	0.23	0.20	-0.16	0.62	0.23	0.14	-0.04	0.49
3rd quantile	1.04***	0.20	0.65	1.44***	0.78***	0.14	0.51	1.05
4th quantile	1.37***	0.21	0.95	1.79***	1.49***	0.15	1.20	1.78
Lowest incomes	1.91***	0.22	1.47	2.35***	1.93***	0.15	1.62	2.23
Attainment grouping at age 7* Family income (Ref. Highest incomes)								
Attainment grouping at age 7 * 2nd quantile	0.22	0.31	-0.39	0.82	0.09	0.21	-0.33	0.51
Attainment grouping at age 7 * 3rd quantile	0.05	0.31	-0.56	0.66	0.09	0.21	-0.33	0.51
Attainment grouping at age 7 * 4th quantile	0.38	0.32	-0.25	1.00	0.17	0.22	-0.26	0.60

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	SE	95% CI	95% CI	B	SE	95% CI	95% CI
Attainment grouping at age 7 * Lowest incomes	0.70*	0.33	0.06	1.33	0.22	0.23	-0.23	0.66
Family income * age								
2nd quintile * age	-0.18	0.11	-0.39	0.04	-0.08	0.07	-0.24	0.04
3rd quintile * age	-0.04	0.11	-0.23	0.44	0.04	0.08	-0.18	0.12
4th quintile * age	-0.10	0.12	-0.16	0.55	0.19	0.08	-0.15	0.15
Lowest incomes * age	0.08	0.12	-0.48	0.24	0.26	0.08	-0.18	0.22
Attainment grouping at age 7 * Highest incomes * age	0.32	0.17	0.00	0.65	0.17	0.11	-0.06	0.39
Attainment grouping at age 7 * 3rd quintile * age	0.10	0.17	-0.23	0.44	0.04	0.12	-0.19	0.27
Attainment grouping at age 7 * 4th quintile * age	0.20	0.18	-0.15	0.55	-0.02	0.12	-0.27	0.22
Attainment grouping at age 7 * Lowest incomes * age	-0.12	0.18	-0.48	0.24	-0.04	0.13	-0.30	0.22
MCS stratum (Ref. England-Advantaged)								

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	SE	95% CI	95% CI	B	SE	95% CI	95% CI
England - Disadvantaged	0.37**	0.14	0.11	0.63	0.35***	0.10	0.16	0.54
England - Ethnic	-0.09	0.21	-0.50	0.33	-0.01	0.16	-0.33	0.32
Wales - Advantaged	0.10	0.27	-0.43	0.63	-0.30	0.17	-0.64	0.04
Wales - Disadvantaged	-0.37	0.20	-0.77	0.02	-0.17	0.14	-0.44	0.11
Scotland – Advantaged	-0.19	0.21	-0.61	0.22	-0.42**	0.15	-0.70	-0.13
Scotland - Disadvantaged	-0.12	0.21	-0.54	0.29	-0.15	0.16	-0.46	0.15
Northern Ireland – Advantaged	-4.25	3.08	-10.28	1.78	-0.73***	0.17	-1.07	-0.39
Northern Ireland - Disadvantaged	1.04	1.28	-1.47	3.55	-0.28	0.16	-0.60	0.05
Female	-1.26***	0.10	-1.46	-1.07	-1.14***	0.06	-1.26	1.01
School type (Ref. State-maintained)								
Fee-paying	-0.90	0.24	-1.38	-0.42	-0.73***	0.20	-1.12	-0.33
Faith	-0.38	0.12	-0.62	-0.15	-0.26**	0.09	-0.43	-0.09
Classroom size	-0.04	0.01	-0.06	-0.02	-0.04***	0.01	-0.06	-0.02
Disruptive peers in CM's class	0.31	0.11	0.09	0.52	0.36***	0.10	0.17	0.55
Number of pupils with SEN in child's class	0.04	0.02	0.00	0.08	0.06****	0.02	0.03	0.10
Number of pupils with EAL in child's class	-0.02	0.01	-0.04	0.00	-0.01	0.01	-0.03	0.00

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Constant	4.83***	0.32	4.20	5.45	4.67***	0.26	4.16	5.18
Random-effects Parameters								
Between-child intercept variance	2.99***	0.05	2.90	3.08	2.98***	0.03	2.92	3.04
Between-child slope variance	0.96***	0.04	0.88	1.05	0.96***	0.03	0.91	1.02
Between-child intercept and slope covariance	-0.26***	0.03	-0.32	-0.20	-0.31***	0.02	-0.35	-0.27
Between-occasion variance	1.80***	0.02	1.76	1.85	1.79***	0.02	1.76	1.82

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05, [NS] Random effect not significant*

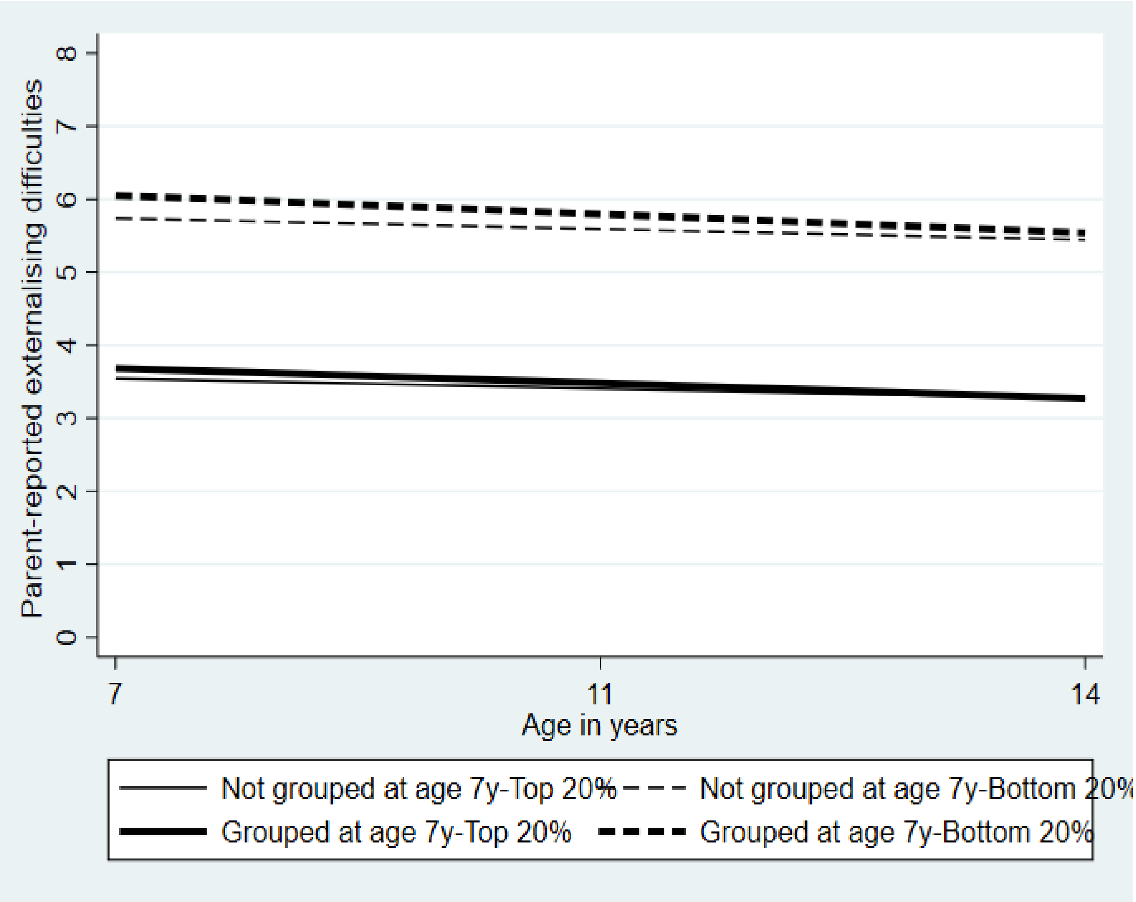

Figure 7. 10: Predicted parent-reported externalising symptoms from age 7 to age14 by family income (Model 3).

7.7.3: Attainment grouping at age 7 and trajectories of externalising symptoms from age 7 to age 14 according to baseline verbal skills

Model 5 (Table 7.10 and Figure 7.11), which tested the interaction between attainment grouping at age 7, baseline verbal skills and age, provided no evidence of effect modification by using both complete cases and multiple imputed data. Similar differences at the level of externalising symptoms were found at age 7 years among those who were grouped by attainment at age 7 years and those who were not, independent of initial verbal skills. In addition, MCS participants who were grouped by attainment at age 7 experienced a similar development in parent-reported externalising symptoms from age 7 to age 14 as children who were not grouped by attainment at age 7, independent of the baseline level of verbal skills (see Figure 7.10).

Table 7. 9: Fixed and random effects for Model 5; the association between attainment grouping at age 7, initial verbal cognitive skills and externalising symptoms

	Model 5-Complete case analysis (n=3,853)				Model 5-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	-0.37	0.53	-1.42	0.68	0.02	0.36	-0.68	0.73
Age	-0.18***	0.03	-0.23	-0.12	-0.20***	0.02	-0.24	-0.16
Attainment grouping at age 7 * age	-0.67*	0.30	-1.25	-0.08	-0.21	0.21	-0.61	0.20
Verbal skills at age 5	-0.07***	0.01	-0.08	-0.05	-0.05***	0.00	-0.06	-0.05
Attainment grouping at age 7*Verbal skills at 5y	0.01	0.01	-0.01	0.03	0.00	0.01	-0.01	0.01
Verbal skills at 5y * age	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00
Attainment grouping at age 7 * verbal skills at age 5 * age	0.01*	0.00	0.00	0.02	0.00	0.00	0.00	0.01
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.60***	0.14	0.33	0.86	0.63***	0.10	0.45	0.82
England - Ethnic	0.19	0.22	-0.23	0.62	0.32*	0.16	0.01	0.63
Wales - Advantaged	0.19	0.27	-0.34	0.73	-0.28	0.17	-0.62	0.06

	Model 5-Complete case analysis (n=3,853)				Model 5-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Wales - Disadvantaged	-0.07	0.20	-0.47	0.33	0.12	0.14	-0.15	0.39
Scotland - Advantaged	-0.15	0.21	-0.57	0.27	-0.41**	0.15	-0.70	-0.12
Scotland - Disadvantaged	0.07	0.22	-0.35	0.49	0.08	0.16	-0.23	0.38
Northern Ireland - Advantaged	-3.47	3.05	-9.45	2.52	-0.33	0.39	-1.10	0.43
Northern Ireland - Disadvantaged	1.36	1.27	-1.12	3.85	0.42	0.43	-0.44	1.28
Female	-1.19***	0.10	-1.39	-0.99	-1.07***	0.07	-1.20	-0.94
School type (Ref. State-maintained)								
Fee-paying	-1.16***	0.24	-1.63	-0.70	-1.15***	0.20	-1.55	-0.75
Faith	-0.40**	0.12	-0.65	-0.16	-0.33**	0.09	-0.50	-0.16
Classroom size	-0.03*	0.01	-0.05	-0.01	-0.04**	0.01	-0.06	-0.02
Disruptive peers in child's class	0.27*	0.11	0.05	0.48	0.33**	0.09	0.14	0.51
Number of pupils with SEN in child's class	0.03	0.02	-0.01	0.07	0.06**	0.02	0.02	0.09
Number of pupils with EAL in child's class	-0.02*	0.01	-0.04	0.00	-0.02***	0.01	-0.04	-0.01
Constant	8.84***	0.46	7.93	9.74	8.48***	0.33	7.84	9.12
Random-effects Parameters								
Between-child intercept variance	2.96	0.05	2.87	3.06	3.00	0.03	2.94	3.06
Between-child slope variance	0.96	0.04	0.88	1.05	0.95	0.03	0.90	1.01

	Model 5-Complete case analysis (n=3,853)				Model 5-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Between-child intercept and slope covariance	-0.26	0.03	-0.32	-0.19	-0.31	0.02	-0.35	-0.27
Between-occasion variance	1.79	0.02	1.74	1.84	1.79	0.02	1.76	1.82

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** p<0.001, ** <0.01, * <0.05, ^{NS} Random effect not significant

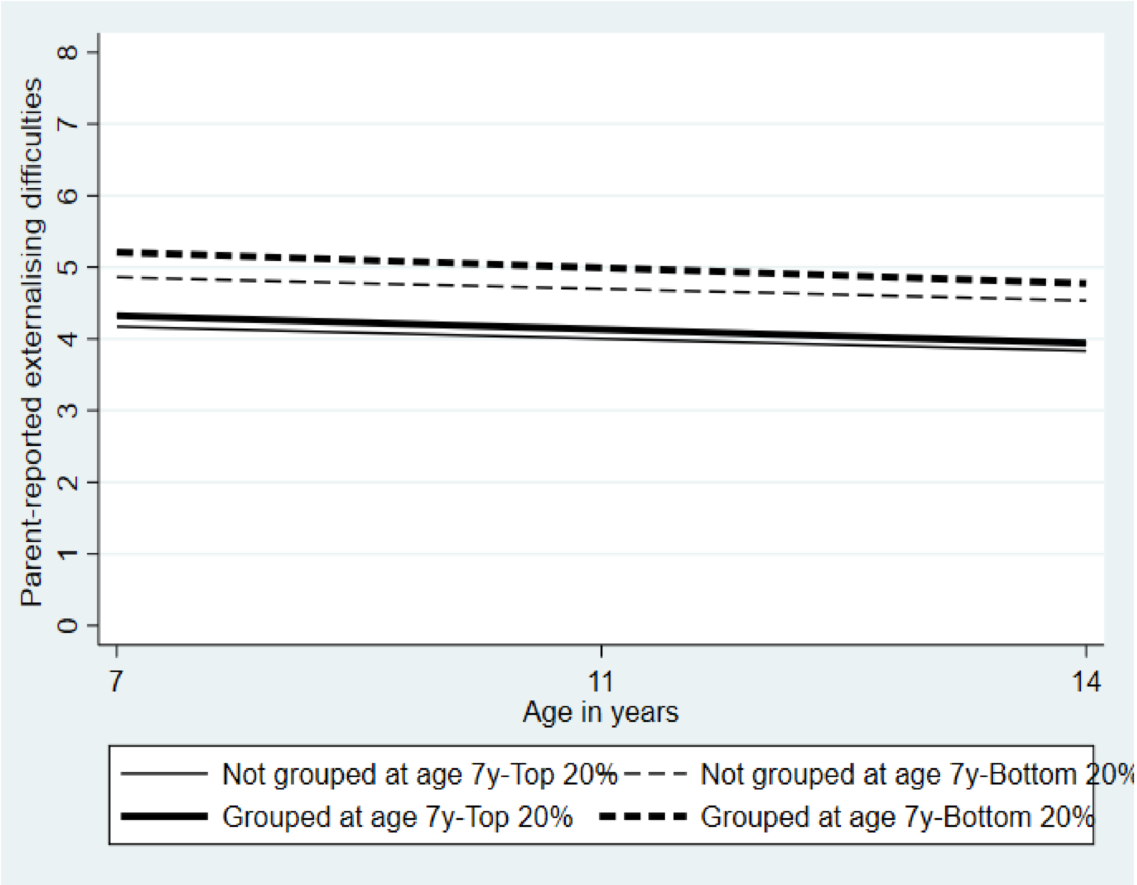

Figure 7. 11 :Attainment grouping at age 7 and trajectories of externalising symptoms from age 7 to age 14 by baseline verbal skills (N=8,629).

7.7.4: Attainment grouping at age 7 and trajectories of internalising symptoms from age 7 to age 14

Model 1 showed a non-significant difference between early attainment grouping at age 7 and children's internalising symptoms at the baseline for age (age 7-intercept) in complete cases (Table 7.12) and multiple imputed data (Table 7.11). A non-significant difference in the rate of change in internalising symptoms between ages 7 and 14 years was also found (see also Figure 7.12).

The random effects coefficients of Model 1 also showed that there was considerable variability in children's parent-reported internalising symptoms at age 7, with these differences increasing with the passage of time.

Adjustment for school-related confounders (Model 2) also suggested the lack of a statistically significant association between early attainment grouping and the level of internalising symptoms. Model 2 highlighted a statistically significant increase in internalising symptoms from age 7 to age 14. Significant associations were also observed between living in disadvantaged and ethnically diverse areas in England and higher internalising symptoms than children living in advantaged areas in England. In contrast, there was a strong association between living in advantaged areas in Scotland and Northern Ireland and lower internalising symptoms. It was also observed that statistical control for school-related confounding attenuated the coefficients representing children who lived in ethnic and disadvantaged areas in England, Wales, Scotland and Northern Ireland, while it strengthened associations between living in advantaged areas in Scotland or Northern Ireland relative to children living in advantaged areas in England and internalising symptoms.

Statistically significant relationships were also found between attending fee-paying and faith primary schools in the UK and fewer internalising symptoms, whereas a bigger classroom size was associated with lower parent-reported internalising symptoms as well. Higher internalising psychological symptoms were associated with having peers with disruptive behaviour and being in a classroom with a greater number of peers with SEN and EAL.

For all parameters of Model 1 and Model 2, the coefficients were similar, whereas the 95% CI were narrower and standard errors were smaller for models using multiple imputed data, demonstrating the efficiency gain and less biased results of multiple imputation over complete cases.

Table 7. 10: Fixed and random effects of the partially (Model 1) and fully adjusted (Model 2) growth curve model p-reported internalising symptoms (multiple imputed data)

	Model 1-Multiple imputation (n=8636, m=50)				Model 2-Multiple imputation (n=8636; m=50)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.07	0.06	-0.05	0.19	0.03	0.06	-0.09	0.14
Age	0.51***	0.03	0.46	0.56	0.50***	0.02	0.46	0.54
Attainment grouping at age 7* age	-0.02	0.04	-0.10	0.06	-0.02	0.04	-0.09	0.06
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.71**	0.08	0.56	0.86	0.58***	0.08	0.43	0.73
England - Ethnic	0.97***	0.11	0.76	1.17	0.69***	0.14	0.41	0.97
Wales - Advantaged	0.03	0.14	-0.24	0.30	-0.18	0.14	-0.46	0.09
Wales - Disadvantaged	0.45***	0.10	0.25	0.65	0.10	0.12	-0.13	0.32
Scotland - Advantaged	-0.22	0.12	-0.45	0.01	-0.43***	0.12	-0.67	-0.20
Scotland - Disadvantaged	0.25*	0.13	0.00	0.50	0.07	0.13	-0.18	0.32
Northern Ireland - Advantaged	-0.25	0.14	-0.52	0.01	-0.50***	0.14	-0.78	-0.22
Northern Ireland - Disadvantaged	0.27*	0.12	0.03	0.52	0.08	0.14	-0.20	0.35
Female	-0.02	0.06	-0.12	0.09	-0.02	0.06	-0.13	0.08
School type (Ref. State-maintained)								
Fee-paying					-1.31***	0.16	-1.62	-1.00
Faith					-0.24**	0.07	-0.38	-0.10

	Model 1-Multiple imputation (n=8636, m=50)				Model 2-Multiple imputation (n=8636; m=50)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Classroom size					-0.05***	0.01	-0.06	-0.03
Disruptive peers in CM's class					0.18*	0.08	0.02	0.35
Number of pupils with SEN in CM's class					0.05***	0.01	0.02	0.08
Number of pupils with EAL in CM's class					0.02*	0.01	0.00	0.03
Constant	2.35***	0.06	2.23	2.46	3.59***	0.21	3.17	4.01
Random-effects Parameters								
Between-child intercept variance	2.11***	0.03	2.05	2.17	2.08***	0.03	2.02	2.14
Between-child slope variance	0.88***	0.03	0.82	0.94	0.88***	0.03	0.82	0.94
Between-child intercept and slope covariance	0.07*	0.04	0.00	0.15	0.07*	0.04	0.00	0.15
Between-occasion variance	1.89***	0.02	1.85	1.93	1.89***	0.02	1.85	1.93

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** p<0.001, ** <0.01, * <0.05, ^{NS} Random effect not significant*

Table 7. 11: Fixed and random effects of the partially (Model 1) and the fully-adjusted (Model 2) multilevel growth curve model of parent-reported internalising symptoms (complete cases)

	Model 1-Complete case analysis (n=3853)				**Model 2-Complete case analysis (n=3853)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.10	0.09	-0.07	0.28	0.06	0.08	-0.11	0.22
Age	0.51***	0.04	0.44	0.59	0.51***	0.03	0.45	0.56
Attainment grouping at age 7* age	-0.02	0.06	-0.13	0.09	-0.02	0.04	-0.14	0.09
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.80***	0.10	0.60	1.00	0.64***	0.11	0.43	0.85
England - Ethnic	0.73***	0.15	0.44	1.01	0.44**	0.17	0.11	0.77
Wales - Advantaged	0.12	0.22	-0.31	0.54	-0.11	0.22	-0.54	0.32
Wales - Disadvantaged	0.41**	0.15	0.11	0.70	0.03	0.16	-0.29	0.34
Scotland - Advantaged	0.03	0.17	-0.30	0.36	-0.22	0.17	-0.55	0.11
Scotland - Disadvantaged	0.35*	0.17	0.01	0.68	0.11	0.17	-0.22	0.44
Northern Ireland - Advantaged	-0.60	2.50	-5.50	4.30	-0.78	2.47	-5.62	4.06
Northern Ireland - Disadvantaged	0.12	1.03	-1.89	2.14	-0.06	1.02	-2.05	1.93
Female	-0.09	0.08	-0.25	0.07	-0.06	0.08	-0.22	0.10
School type (Ref. State-maintained)								
Fee-paying					-1.41***	0.19	-1.78	-1.04
Faith					-0.31**	0.10	-0.50	-0.12

	Model 1-Complete case analysis (n=3853)				Model 2-Complete case analysis (n=3853)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Classroom size					-0.05***	0.01	-0.06	-0.03
Disruptive peers in CM's class					0.16	0.09	-0.01	0.33
Number of pupils with SEN in CM's class					0.05**	0.02	0.02	0.08
Number of pupils with EAL in CM's class					0.02*	0.01	0.00	0.03
Constant	2.32***	0.08	2.16	2.47	3.64***	0.23	3.19	4.10
Random-effects Parameters								
Between-child intercept variance	2.06***	0.05	1.97	2.15	2.02***	0.05	1.93	2.11
Between-child slope variance	0.90***	0.05	0.81	0.99	0.90***	0.05	0.81	1.00
Between-child intercept and slope covariance	0.09^{NS}	0.06	-0.02	0.20	0.09^{NS}	0.06	-0.02	0.20
Between-occasion variance	1.91***	0.02	1.86	1.96	1.91***	0.03	1.86	1.96

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05,* ^{NS} *Random effect not significant*

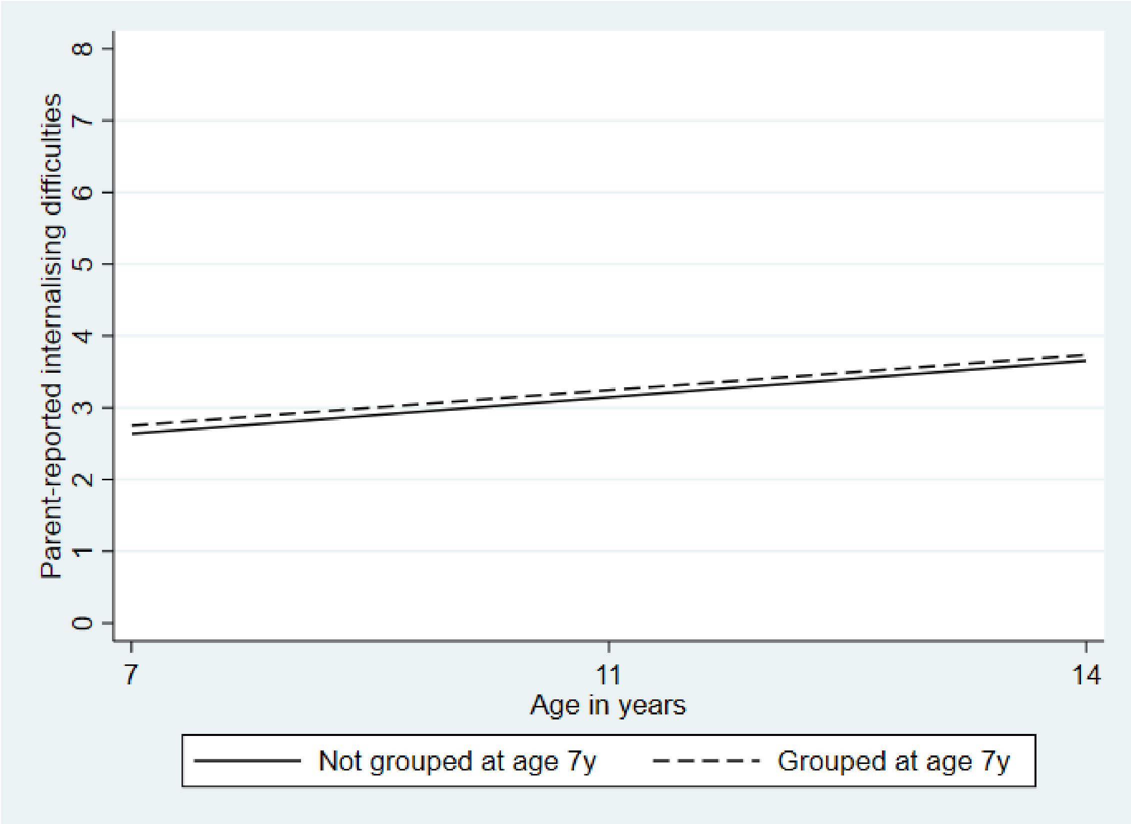

Figure 7. 12: Attainment grouping at age 7 and symptoms of internalising difficulties from 7 to 14 years of age (N=8,629)

7.7.5: Attainment grouping at age 7 and trajectories of internalising symptoms from 7 to 14 years according to family income

Model 4 which tested the interaction between attainment grouping at age 7 and family income over time, showed no evidence of effect modification in both complete cases (Table 7.14) and multiple imputed data (Table 7.13 and Figure 7.13).

Table 7. 12: Fixed and random effects for Model 4; the association between attainment grouping at age 7, family income and parent-reported internalising symptoms

	Model 4-Comlete case analysis (N=3,853)				**Model 4-Multiple imputation (n=8,629)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.17	0.17	-0.17	0.53	0.08	0.13	-0.18	0.34
Age	0.51***	0.03	0.45	0.57	0.47***	0.05	0.37	0.54
Attainment grouping at age 7 * age					-0.05	0.08	-0.21	0.11
Family income (Ref. Highest incomes								
2nd quantile	0.42**	0.16	0.11	0.73	0.35**	0.11	0.14	0.57
3rd quantile	0.77***	0.16	0.46	1.08	0.58***	0.11	0.36	0.80
4th quantile	1.17***	0.17	0.84	1.50	1.29***	0.12	1.06	1.52
Lowest incomes	1.40***	0.18	1.05	1.74	1.44***	0.12	1.19	1.68
Family income (Ref. Highest incomes* age								
2nd quantile *age	0.04	0.11	-0.18	0.25	0.03	0.07	-0.11	0.17
3rd quantile * age	-0.10	0.11	-0.32	0.11	-0.02	0.07	-0.17	0.12
4th quantile *age	0.08	0.12	-0.15	0.31	0.14	0.08	-0.02	0.30
Lowest incomes * age	0.03	0.12	-0.21	0.27	0.04	0.08	-0.31	0.47

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	SE	95% CI	95% CI	B	SE	95% CI	95% CI
Attainment grouping at age 7* Family income (Ref. Highest incomes)								
Attainment grouping at age 7 * 2nd quantile	-0.05	0.24	-0.53	0.42	0.01	0.17	-0.33	0.35
Attainment grouping at age 7 * 3rd quantile	-0.24	0.24	-0.72	0.23	-0.06	0.17	-0.40	0.27
Attainment grouping at age 7 * 4th quantile	-0.11	0.25	-0.60	0.38	-0.16	0.18	-0.51	0.19
Attainment grouping at age 7 * Lowest incomes	0.40	0.26	-0.10	0.91	0.04	0.18	-0.32	0.40
Attainment grouping at age 7* Family income (Ref. Highest incomes) * age								
Attainment grouping at age 7 * 2nd quantile *age	0.17	0.17	-0.16	0.51	0.06	0.11	-0.16	0.29
Attainment grouping at age 7 * 3rd quantile*age	0.43*	0.17	0.09	-0.77	0.23	0.12	0.00	0.46
Attainment grouping at age 7 * 4th quantile * age	0.09	0.18	-0.26	0.45	-0.06	0.13	-0.32	0.20

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Attainment grouping at age 7 * Lowest incomes *age	0.09	0.18	-0.27	0.46	-0.07	0.13	-0.33	0.19
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.39***	0.11	0.19	0.60	0.30***	0.08	0.15	0.45
England - Ethnic	0.14	0.17	-0.18	0.47	0.32*	0.14	0.05	0.60
Wales - Advantaged	-0.11	0.21	-0.52	0.31	-0.20	0.14	-0.47	0.07
Wales - Disadvantaged	-0.19	0.16	-0.49	0.12	-0.17	0.11	-0.39	0.06
Scotland - Advantaged	-0.15	0.17	-0.48	0.17	-0.36**	0.12	-0.60	-0.13
Scotland - Disadvantaged	-0.06	0.17	-0.39	0.26	-0.14	0.13	-0.38	0.11
Northern Ireland - Advantaged	-1.60	2.42	-6.34	3.15	-0.55***	0.14	-0.81	-0.28
Northern Ireland - Disadvantaged	-0.51	0.99	-2.46	1.45	-0.28*	0.13	-0.54	-0.02
Female	-0.11	0.08	-0.27	0.05				
School type (Ref. State-maintained)								
Fee-paying	-0.87***	0.19	-1.249	-0.50	-0.79****	0.16	-1.10	-0.47
Faith	-0.24*	0.10	-0.423	-0.05	-0.16*	0.07	-0.30	-0.03
Classroom size	-0.04***	0.01	-0.051	-0.02	-0.04***	0.01	-0.05	-0.02
Disruptive peers in CM's class	0.13	0.09	-0.034	0.30	0.17*	0.08	0.01	0.33
Number of pupils with SEN in child’s class	0.03*	0.02	0.001	0.06	0.04**	0.01	0.02	0.07

	Model 4-Comlete case analysis (N=3,853)				Model 4-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Number of pupils with EAL in child's class	0.00	0.01	-0.012	0.02	0.00	0.01	-0.01	0.02
Constant	2.81***	0.25	2.320	3.31	2.82***	0.22	2.38	3.26
Random-effects Parameters								
Between-child intercept variance	1.96***	0.05	1.872	2.05	2.02***	0.03	1.96	2.08
Between-child slope variance	0.90***	0.05	0.810	1.00	0.88***	0.03	0.82	0.94
Between-child intercept and slope covariance	0.09^{NS}	0.06	-0.027	0.20	0.07	0.04	0.00	0.15
Between-occasion variance	1.91***	0.03	1.865	1.96	1.89***	0.02	1.85	1.93

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05,* ^{NS} *Random effect not significant*

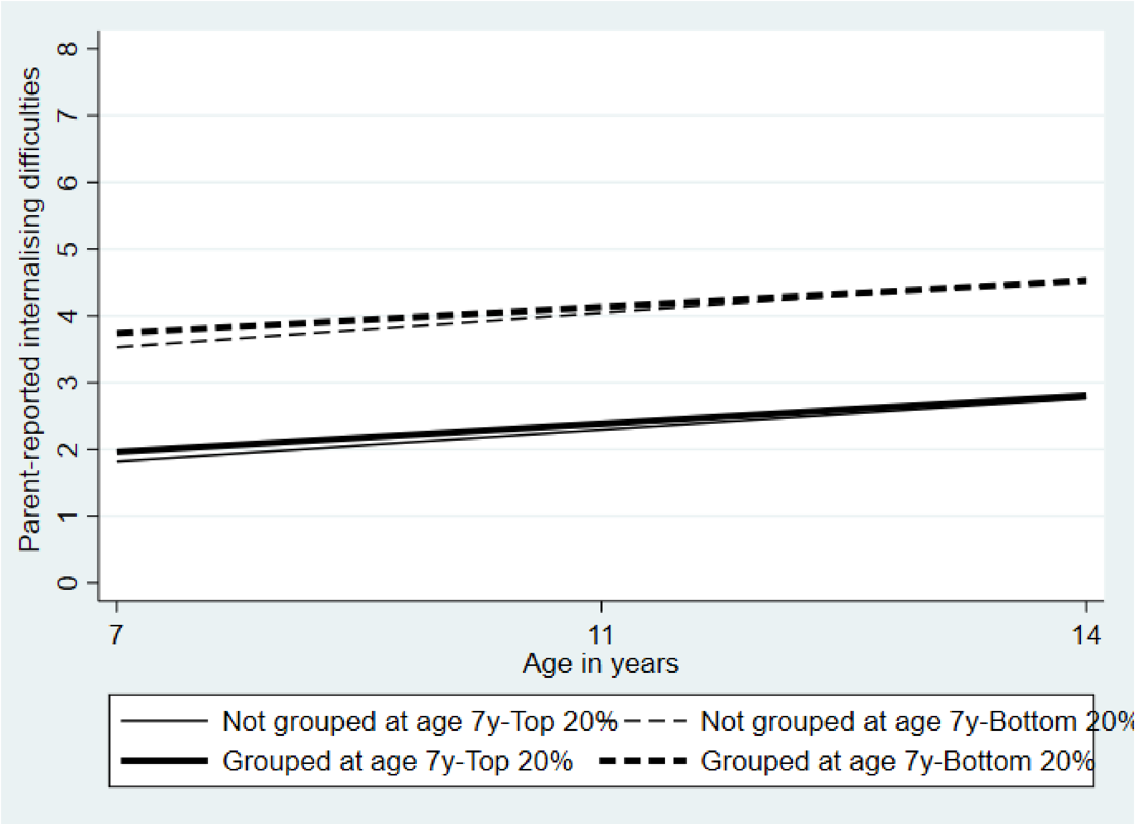

Figure 7. 13: Attainment grouping at age 7 and symptoms of internalising difficulties from 7 to 14 years of age by family income (N=8,629)

7.7.6: Attainment grouping at age 7 and trajectories of internalising symptoms from age 7 to age 14 according to baseline verbal skills

Model 5 which tested the interaction between attainment grouping at age 7, initial verbal skills and age showed evidence of effect modification when analysis was conducted using multiple imputed data (Table 7.14). At age 7, the difference in internalising symptoms between high-attaining and low-attaining MCS participants at primary school entry was slightly higher among MCS participants who were grouped by attainment at age 7 (see Figure 7.14). This difference significantly decreased over time because of a slightly faster increase in internalising symptoms among MCS participants who entered UK primary schools with a higher level of verbal skills, and they were grouped by attainment at age 7.

Table 7. 13: Fixed and random effects for Model 5; the association between attainment grouping at age 7, initial verbal cognitive skills and parent-reported internalising symptoms from age 7 to age 11

	Model 5- Complete case analysis (n=3,658)				**Model 5-Multiple imputation (n=8,629)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.60	0.46	-0.30	1.51	0.71*	0.32	0.08	1.35
Age	0.83***	0.20	0.43	1.23	0.74***	0.14	0.47	1.01
Attainment grouping at age 7 * age	-0.46	0.30	-1.06	0.14	-0.57**	0.21	-0.99	-0.15
Verbal skills at age 5	-0.03***	0.01	-0.04	-0.02	-0.02***	0.00	-0.03	-0.02
Verbal skills at age 5 * age	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00
Attainment grouping at age 7*Verbal skills at 5y	-0.01	0.01	0.00	0.01	-0.01*	0.01	-0.02	0.00
Attainment grouping at age 7 * verbal skills at age 5 * age	0.01	0.00	0.00	0.02	0.01**	0.00	0.00	0.02
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.56***	0.11	0.35	0.77	0.51***	0.08	0.36	0.66
England - Ethnic	0.36*	0.17	0.02	0.70	0.51***	0.13	0.25	0.76
Wales - Advantaged	-0.09	0.22	-0.51	0.33	-0.21	0.14	-0.48	0.07
Wales - Disadvantaged	0.07	0.16	-0.25	0.38	0.01	0.11	-0.21	0.23

	Model 5- Complete case analysis (n=3,658)				Model 5-Multiple imputation (n=8,629)			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Scotland - Advantaged	-0.16	0.17	-0.49	0.18	-0.37**	0.12	-0.61	-0.14
Scotland - Disadvantaged	0.07	0.17	-0.27	0.40	0.05	0.13	-0.20	0.30
Northern Ireland - Advantaged	-0.99	2.42	-5.74	3.76	-0.24	0.38	-1.00	0.52
Northern Ireland - Disadvantaged	-0.23	0.99	-2.19	1.72	0.31	0.43	-0.55	1.16
Female	0.00	0.08	-0.16	0.16	0.07	0.06	-0.04	0.18
School type (Ref. State-maintained)								
Fee-paying	-1.12***	0.19	-1.49	-0.75	-1.15***	0.16	-1.46	-0.84
Faith	-0.24*	0.10	-0.43	-0.05	-0.23**	0.07	-0.37	-0.09
Classroom size	-0.03**	0.01	-0.04	-0.01	-0.04****	0.01	-0.05	-0.03
Disruptive peers in CM's class	0.11	0.09	-0.06	0.28	0.16	0.08	0.00	0.32
Number of pupils with SEN in child's class	0.03	0.02	-0.01	0.06	0.05**	0.02	0.02	0.08
Number of pupils with EAL in child's class	0.00	0.01	-0.01	0.02	0.01	0.01	-0.01	0.02
Constant	4.83***	0.37	4.11	5.55	5.06***	0.28	4.51	5.61
Random-effects Parameters								
Between-child intercept variance	1.98	0.05	1.89	2.08	2.06	0.03	2.00	2.13
Between-child slope variance	0.91	0.05	0.82	1.00	0.89	0.04	0.82	0.96

	Model 5- Complete case analysis (n=3,658)				**Model 5-Multiple imputation (n=8,629)**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Between-child intercept and slope covariance	0.07	0.06	-0.04	0.19	0.06	0.04	-0.02	0.14
Between-occasion variance	1.89	0.03	1.84	1.94	1.88	0.02	1.85	1.92

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05, NS Random effect not significant*

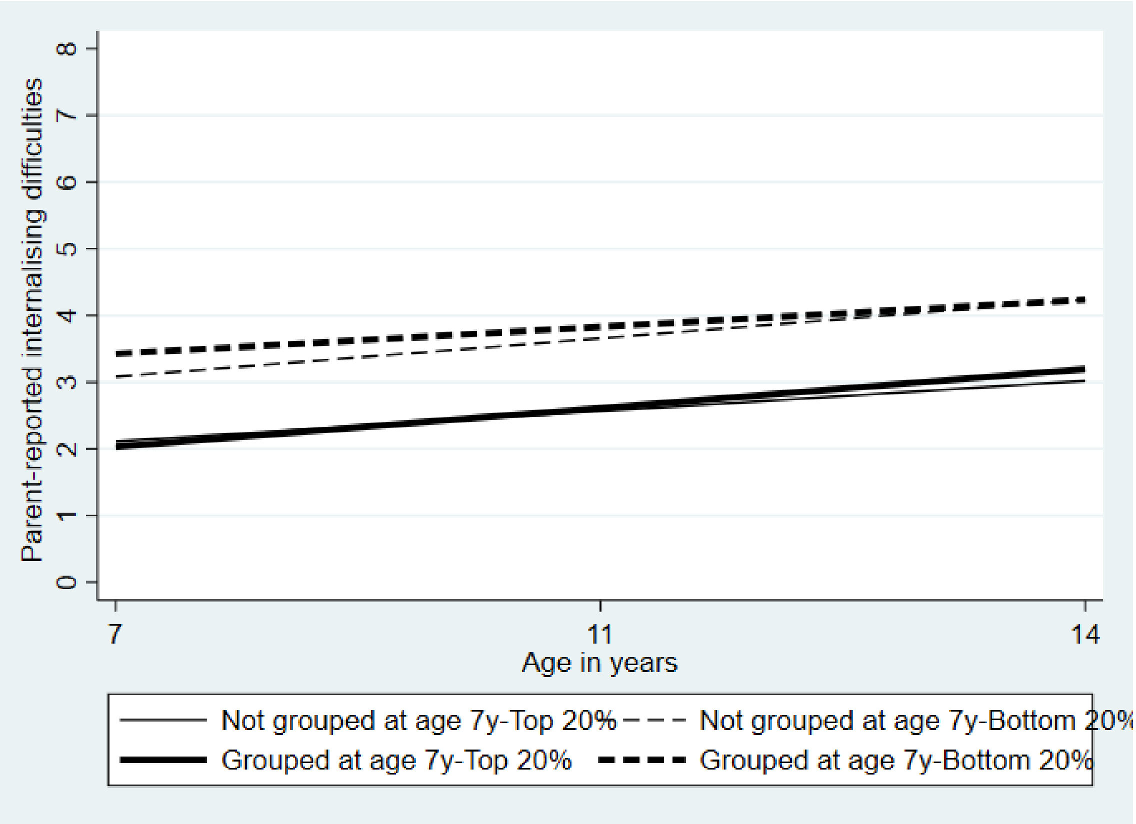

Figure 7. 14: Attainment grouping at age 7 and internalising symptoms from age 7 to age 14 by baseline verbal skills (N=8,629)

7.8: Sensitivity analysis

7.8.1: Excluding the Scottish and the Northern Irish MCS sample

Education is a devolved issue in the UK. Although England, Wales, Scotland, and Northern Ireland follow a broadly similar structure, there are significant variations between the primary school systems, ranging from the age that children start school to the time and type of tests (SATs) they take.

In England and Wales children start primary school before turning five. In Scotland children with birthdays between March and August start school in August following their fifth birthday, while children with birthdays between September and February can start school before the fifth birthday (this can be deferred until the following September). Northern Ireland has the lowest compulsory school starting age in Europe: children who have reached the age of four on July 1st can start school the following September.

School years are grouped together in larger stages or phases, with different requirements and outcomes for each. In England, the Foundation Stage includes preschool, nursery, and Reception. In Wales, the Foundation Phase covers children from the age of 3 to age 5 years. In both England and Wales, KS1 cover Years 1 and Year 2, while KS2 cover Years 3 to 6. In Northern Ireland, the Foundation Phase covers Years 1 and 2, KS1 Years 3 and 4, and KS 2 Years 5 to 7. In Scotland, there are no phases or stages; the curriculum for excellence is followed from age 3 to 18.

Considering these differences, the adoption and the intensity of attainment grouping in the early years of primary school years may vary between the four UK countries, and thus the relationship between attainment grouping at age 7 and trajectories of children's psychological symptoms may be different for each UK country. For this reason, multilevel multivariable regression analyses were re-applied in England and Wales only (the education systems of the two countries are very similar) to explore whether the strength of the association between attainment grouping at KS1 and trajectories of MCS participants' psychological symptoms up to KS 3 (age 14) was lessened by the inclusion of the Scottish and Northern Irish MCS sample.

Table A4.2 to A4.4 highlight very similar estimates compared to those obtained when the MCS samples of the four UK countries were considered.

7.9: Summary of findings

A summary of the findings of this chapter's analyses addressing objectives 3.1 to 3.3 and their related hypotheses is presented below.

Objective 3.1: **To examine whether attainment grouping at age 7 is associated with a less favourable change in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.**

Being exposed to attainment grouping at age 7 was not associated with a slower decrease in parent-reported externalising and a faster increase in parent-reported internalising psychological symptoms between ages 7 and 14 years.

Objective 3.2: **To test whether attainment grouping at age 7 is related to an increase in socioeconomic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.**

Attainment grouping at age 7 was not associated with an increase in socio-economic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.

Objective 3.3: **To test whether attainment grouping at age 7 is associated with an increase in differences in externalising and internalising psychological symptoms between 7 and 14 years among MCS participants with higher and lower levels of verbal skills at age 5.**

The gap in externalising and internalising psychological symptoms did not increase between age 7 and 14, contingent on differences in verbal skills at age 5 for those MCS participants in attainment groups at age 7. In contrast, a slight decrease in the difference in externalising psychological symptoms was observed among MCS participants with different levels of verbal skills at age 5, and they were grouped by attainment at age 7.

7.10: Discussion

Using data from the MCS has enabled this research to demonstrate with a relatively large, representative sample of children from a range of different school environments in the UK that children who were taught in streams or sets at age 7 experienced a similar rate of change in terms of their externalising or internalising symptoms to children who were not allocated to streams or sets at age 7. No evidence was found that the association between attainment grouping at age 7 and the rate of change in externalising or internalising symptoms was different according to family income. In a similar vein, the rate of externalising symptoms did not differ between children who were grouped by attainment at age 7 and those who were not, regardless of baseline verbal skills (age 5). Nevertheless, it was found that the difference in internalising symptoms between children who started primary school with higher and lower levels of verbal skills and who were grouped by attainment at age 7 became smaller over time. This is because internalising symptoms grew slightly faster for high-attaining children who were grouped by attainment at age 7 compared to low attaining children who were also grouped by attainment at age 7.

7.10.1: Comparison to previous literature

The overall finding of no significant association between attainment grouping at age 7 and trajectories of externalising or internalising symptoms from age 7 to age 14 complements the finding reported in chapter 6 of no effect of attainment grouping transition on parent-reported or teacher-reported externalising or internalising symptoms. Therefore, it could be argued that in general attainment grouping during primary school, whether it happens at age 7 or at age 11 seems to play little or no role for the development of externalising or internalising symptoms during childhood or early-adolescence. Overall, these findings do not support recently raised concerns from teachers in English primary schools that early attainment grouping could have a potentially negative effects upon some children’s mental health (Bradbury and Roberts-Holmes, 2017). Further research should be undertaken to investigate whether early attainment grouping may affect some children’s mental health differently based on their ethnic group status, whether they have SEN, or they speak English as an additional language.

7.10.2: Strengths and limitations

The main strength of this chapter’s analyses is the use longitudinal data in a large, population-based sample. This was important in order to examine the association between attainment grouping at age 7 and children’s trajectories of externalising or internalising psychological symptoms over a longer study period spanning the important transition to puberty and secondary school, thus avoiding reverse causality bias.

The findings in this chapter are subject to at least two limitations. First, some MCS participants might have been taught in attainment groups for several curriculum subjects, while others might have had minimal attainment grouping exposure at age 7 (e.g one curriculum subject). It may be then possible then that a dose-response relationship exists between the number of attainment groups at age 7 and different rates of development in externalising and internalising psychological symptoms from 7 to 14 years.

Second, an issue that was not addressed in this study was the role of secondary school. Information on attainment grouping during secondary school is lacking in the MCS. It may possible that trajectories of externalising and internalising psychological symptoms are different for children who are grouped by attainment in both primary and secondary school compared to children who are grouped by attainment only in primary school.

Third with only three time points of data on externalising and internalising psychological symptoms , the possibilities for modelling the functional form of children's individual trajectories were limited.

Chapter 8: Discussion and Conclusion

8.1: Chapter overview

This chapter will summarise the results of this PhD's empirical chapters. Chapter 5 explored the association between transition to attainment grouping between ages 7 and 11 years (KS 2) and the change in MCS participants verbal skills from age 7 to age11 by using a fixed effects difference in differences design. It was also explored whether the transition to attainment grouping was associated with an increase in socio-economic inequalities in children's verbal skills or whether the verbal skills gap at school entry (high versus low scores) widened for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups in English and Welsh primary school.

Chapter 6 explored the association between the transition to attainment grouping between ages 7 and 11 years and the change in MCS participants' parent- and teacher-reported externalising and internalising psychological symptoms from age 7 to age 11 in English and Welsh primary school. Further, it was explored whether the transition to attainment grouping was associated with an increase in socio-economic inequalities in children's psychological symptoms. Last but not least, it was explored whether the transition from mixed grouping to attainment grouping between ages 7 and 11 years was associated with an increased gap in externalising and internalising psychological symptoms between MCS participants with higher verbal skills and participants with lower verbal skills at primary school entry.

Chapter 7 explored the long-term association between attainment grouping at age 7 and trajectories of parent-reported externalising and internalising psychological symptoms from age 7 to age 14 in UK primary school, while effect modification by family income and verbal skills at age 5 was tested.

Further, this chapter discusses the findings of this PhD's in relation to the objectives and hypotheses described in chapter 3. Then it goes on discussing possible policy implications of the findings, the strengths, and the limitations of the current research, and it concludes on discussing the directions for future research on the effects of attainment grouping.

8.2: Summary of findings

Objective 1.1 **To examine whether a transition from mixed attainment grouping to an attainment grouping (streaming/setting for literacy) between ages 7 and 11 years is related to progress in children's verbal skills.**

No overall effect of the transition to attainment grouping between ages 7 and 11 years on MCS participants' verbal skills from age 7 to age 11 was observed. Therefore, the hypothesis 1.1 of a negative effect of the transition to attainment grouping on children's verbal skills improvement was not supported. This finding is consistent with previous literature of no overall effect of attainment grouping on standardised achievement tests in the UK and the US (Ireson et al., 2005, Steenbergen-Hu et al., 2016). Evidence from this PhD and earlier research demonstrates that attainment grouping in primary schools may not itself raise attainment for all children.

Objective 1.2 **To test whether a transition from mixed to attainment grouping between ages 7 and 11 years is related to a change in socioeconomic inequalities in children's verbal skills between ages 7 and 11 years.**

Little evidence was provided of an increase in socioeconomic inequalities in children's verbal skills from age 7 to age 11 in response to attainment grouping transition between ages 7 and 11 years. Therefore, compelling evidence for the hypothesis 1.2; of an increase in socioeconomic gaps in children's verbal skills due the transition to attainment grouping was not be provided.

Objective 1.3 **To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a change in the gap in verbal skills between age 7 and 11 contingent on the gap in verbal skills at age 5.**

The gap in verbal skills at school entry (high versus low scores) widened for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups. Thus, hypothesis 1.3 was supported.

This finding lends support to the divergence hypothesis during primary education in England and Wales. The divergence hypothesis posits that studying and learning in mixed-attainment classrooms would prevent the formation of a gap between high and low-attaining students greater than expected given the initial differences between them.

This finding also suggests that attainment grouping may undermine the efforts of the UK government to raise attainment for all children in England and Wales whatever their readiness for primary school. Attainment grouping is not statutory for UK schools which are free to choose what is best for their students (Department for Education., 2010) Given that attainment grouping may be counterproductive in reducing the attainment gap, primary schools in England and Wales may also need to take this into account in their attempt to raise academic standards for their students.

Objective 2.1: **To examine whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with a less favourable change in children's externalising or internalising psychological symptoms from age 7 to age 11.**

No evidence was found in support of hypothesis 2.1. The transition to attainment grouping between ages 7 and 11 years was not associated with a less favourable change in either parent-reported nor teacher-reported externalising and internalising psychological symptoms from age 7 to age 11. Being exposed to a transition from mixed grouping to attainment grouping between ages 7 and 11 years is not associated with a smaller decrease in parent- or teacher-reported externalising psychological symptoms and with a bigger increase in parent- or teacher-reported internalising psychological symptoms from age 7 to age 11.

This finding could provide some comfort for parents', teachers', and policy makers' concerns on possible negative consequences of attainment grouping on students' externalising and internalising psychological symptoms.

Objective 2.2: **To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is related to an increase in socioeconomic inequalities in children's externalising and internalising psychological symptoms between ages 7 and 11 years.**

No evidence was found of an increase in parent-reported and teacher-reported externalising and internalising psychological symptoms from age 7 to age 11 between the most and the least socioeconomically advantaged MCS participants who changed to attainment grouping between age 7 and age 11 and those who did not. Therefore, findings did not support the hypothesis 2.2; an increase in socioeconomic inequalities in children's externalising and internalising psychological symptoms from age 7 to age 11 would be observed.

Objective 2.3: **To test whether a transition from mixed grouping to attainment grouping between ages 7 and 11 years is associated with an increased gap in externalising and internalising psychological symptoms between MCS participants with higher verbal skills and participants with lower verbal skills at age 5.**

The gap in externalising and internalising psychological symptoms for MCS participants with different levels of verbal skills at school entry (high versus low scores) did not widen among those who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups. Thus, hypothesis 2.3 was not supported.

Objective 3.1: **To examine whether attainment grouping at age 7 is associated with a less favourable change in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.**

No evidence was found in support of hypothesis 3.1. Being exposed to attainment grouping at age 7 was not associated with a slower decrease in parent-reported externalising or a faster increase in parent-reported internalising psychological symptoms between ages 7 and 14 years.

This finding complements the finding of chapter 6, where no significant association between the transition to attainment grouping between ages 7 and 11 years and the change parent-reported and teacher-reported externalising and internalising psychological symptoms from age 7 to age 11 was found.

Objective 3.2: **To test whether attainment grouping at age 7 is related to an increase in socioeconomic inequalities in parent-reported externalising and internalising psychological symptoms between 7 and 14 years.**

Attainment grouping at age 7 was not significantly associated with a higher level of parent-reported externalising or internalising symptoms at age 7 among the most and the least socioeconomically advantaged MCS participants, after statistical adjustment for baseline school-level confounders. In addition, it was found that socioeconomic differences in parent- and teacher-reported externalising and internalising psychological symptoms did not significantly change from 7 to 14 years. Evidence was therefore provided against the hypothesis 3.2 of greater socioeconomic differences in MCS participants' psychological symptoms at age 7 and from 7 to 14 years in relation to attainment grouping at age 7.

Objective 3.3: **To test whether attainment grouping at age 7 is associated with an increase in differences in externalising and internalising psychological symptoms between 7 to 14 years among MCS participants) with higher and lower levels of verbal skills at age 5.**

The gap in externalising symptoms did not increase between age 7 and 14, contingent on differences in verbal skills at age 5 for those MCS participants in attainment groups at age 7. Therefore hypothesis 3.3 was not supported for externalising psychological symptoms.

The gap in internalising psychological symptoms between MCS participants in attainment groups at age 7 slightly decreased between age 7 and age 14. MCS participants who started primary school with a lower level of verbal skills and they were taught in attainment groups at age 7 experienced a slightly slower increase in their internalising psychological symptoms that MCS participants who started school with a higher level of verbal skills and they were taught in attainment groups at age 7.

8.3: Policy implications

This is the first UK contemporary study to report that attainment grouping in primary school, on average, is not associated with an increase in verbal cognitive attainment for all children and it is associated with a in widening of the verbal cognitive attainment gap during primary school in England and Wales among children with a low and high level at age 5. The present findings of this PhD complement a limited but growing body of research indicating that attainment grouping in secondary school neither harms students' academic progress, nor has academic benefits, and if anything, it is associated with a widening of attainment gaps (Education Endowment Foundation, 2018a).

Children growing up in poor families emerge from UK schools with substantially lower levels of educational attainment. These educational deficits emerge early in children's lives, even before entry into school, and widen throughout childhood (Crawford et al., 2017, Clifton and Cook, 2012). Early childhood home environment is important in establishing the foundations of optimal learning. What parents do or not do matters for child development and how children will later masterfully navigate life challenges. Socio-economic gradients are apparent early in a child's life in several overlapping domains of the home environment such as the frequency of home learning activities (e.g the frequency of parent reading to the child, visits to the library etc.), family routines such as regular bedtimes and mealtimes, and markers of psychosocial environment such as maternal psychological distress, discipline practices and parent-child relationship (Kelly et al., 2011, Dearden et al., 2011). However, access to early education is partly responsible for the unequal start in primary education. In the UK children from disadvantaged backgrounds are almost twice as likely not to access their funded place for the full five terms as children from more advantaged backgrounds, while children from families where all parents are in paid work are entitled for 30 hours of free early education rather than 15 funded hours (Campbell et al., 2019). Therefore, early childhood home environment and early education experiences could account for the educational differences observed in the beginning of compulsory education.

Attainment grouping could also be one of the possible factors behind this increase in educational differences during primary and secondary school. Thereby, attainment grouping during primary and secondary school in the UK is likely to undermine government attempts to reduce educational inequality and increase social mobility.

Abolishing between-class attainment grouping in the UK is difficult given the current circumstances. The "*Best Practice in Grouping Students"* research project highlighted challenges in recruiting and retaining secondary schools committed in practising mixed attainment grouping (only 18 schools out of 488 secondary schools in England agreed to be randomly allocated to either receive the training in mixed attainment grouping or to be a control group in this research project (Palak et al., 2018)). Given the reluctance of schools to adopt mixed attainment methods and an increased level of parental choice , attention could therefore be focused on applying more equitable principles in attainment grouping with the aim of improving the quality of pedagogy for children with lower prior cognitive or academic attainment and help reduce inequality within school. If the UK primary schools continue to allocate more experienced teachers in charge of students who are doing well due to external pressures such as league tables (Slavin, 1990), instead of those students who struggle academically, or teachers continue to have differing expectations and adopt less challenging pedagogical practices for low-attaining students (Sharples et al., 2011, Boaler et al., 2000), the primary education system is likely to further perpetuate educational inequalities.

The Best Practice in Grouping Students' research project made a series of recommendations in improving existing practices in attainment grouping with regard to efficacy and equity including allocating students in attainment groups solely according to their prior achievement tests rather than their individual or family characteristics, equal access for all students to a rich curriculum, applying high expectations to all attainment groups, regularly retesting and moving students between groups and using a lottery system when assigning borderline students to attainment groups (Francis et al., 2018). So, the issue becomes not so much whether schools use attainment grouping, but how they use it. The best practice in setting intervention which randomised 127 secondary schools in England to either receive the Best Practice in Setting intervention (in English and/or maths) or to continue with their existing setting practices, found no evidence that the intervention improves maths or English attainment for children in Years 7 and 8. However, this finding was dubious given the level of uncertainty and the large amount of attrition (Palak et al., 2018). By focusing on what schools do with attainment grouping, OECD analysts pointed, for example, to Canadian schools that use attainment grouping as much as US schools do, but have more equal achievement outcomes because of how they use attainment groups (OECD, 2011).

This PhD provided evidence of no association between attainment grouping in primary school and externalising psychological symptoms during childhood and adolescence regardless of their initial level of verbal skills. This finding confronts the idea that differentiating learning environments for high-attaining and low-attaining students in primary school reduces misbehaviour by reducing the possibility of distraction from both ends of the attainment distribution (House of Commons., 2011).

8.4: Strengths of the current research

8.4.1: The Millennium Cohort Study

The analyses of this PhD have been benefited by using a large, prospective, nationally representative cohort study, the Millennium Cohort Study. The prospective survey data used allowed the application of longitudinal methods for exploring the association between the change to attainment grouping between ages 7 and 11 years in England and Wales and the change in children's verbal skills and psychological symptoms from age 7 to age 11. It was possible to investigate the long-term association between attainment grouping at age 7 in UK primary schools and trajectories of psychological symptoms from age 7 to age 14.

8.4.2: Adjustment for confounding

The findings of this PhD are an important addition to the current literature as academic research on the consequences of attainment grouping during primary school in a contemporary UK context is currently lacking. The statistical analyses of chapters 5 and 6 attempted to overcome methodological issues of previous academic research such as unobserved time-invariant and time-varying confounding bias. The fixed effects difference-in-difference modelling applied to the statistical analyses of these two chapters accounted for all possible characteristics of the individuals – even without measuring them- so long as those characteristics do not change over time (Allison, 2009).

8.4.3: Multiple Imputation

Missing data on attainment grouping in primary school and the outcomes of interest during childhood and adolescence were imputed by taking advantage of the rich information of the MCS. Comparisons of models using complete cases and multiple imputed data were carried out. This showed that there were no substantive differences between models using complete cases and imputed data. Nevertheless, there was a considerable decrease in the coefficients' standard error in models using multiple imputed data. Models using multiple imputed data were therefore sufficiently robust for the scope of these analyses and more efficient than models using complete cases.

8.4.4: Statistical power of time interactions

The large sample size of the MCS in conjunction with the correction for item non-response (multiple imputation) implies adequate statistical power (> 80%) for analyses of change in verbal skills and psychological symptoms over time in relation to early attainment grouping or the transition to attainment grouping during primary school. The MCS has also repeated measurements of attainment grouping, verbal skills and psychological symptoms over time with increased within-subject correlations (>50%) (Heo and Leon, 2010), which may further increase the accuracy of measuring the true underlying value and may thereby increase the statistical power for these measurements (McConnell and Vera-Hernández, 2015, Gallop et al., 2015)

8.5: Limitations of the current research

Apart from strengths, this research has several limitations that are important to highlight. Some of these limitations have been discussed in detail in each empirical chapter. However, there are other, more general, limitations to be mentioned to situate the findings of this PhD more broadly. Highlighting the limitations of PhD is also useful to point to potential areas for future research (discussed in the next section).

8.5.1: Limits of representativeness

The MCS was sampled by means of a *'disproportionately stratified cluster sample'* of births in the first year of the 21st century. The MCS4 and MCS5 teacher survey questionnaires represent teachers in schools eventually attended by the MCS participants who were selected through this initial sampling procedure, rather being sampled from and directly representing all schools in the four UK countries. The MCS teacher surveys were then dependent on the eventual distribution of MCS participants across participating primary schools at age 7 and age 11 after sampling. Little is also known about the characteristics of teachers and potentially schools who responded to MCS4 and MCS5 surveys and those who did not.

8.5.2: Identification issues

8.5.2.1: Co-occurrence of between-class and within-class attainment grouping

Primary schools in the UK that choose to employ streaming or setting might also teach their students in groups within their class. It is also possible that schools that do not choose to employ between-class attainment grouping (streaming or setting) might group their students within their class. Hence, the estimates of this PhD are not entirely capturing the effect of between-class attainment grouping versus the counterfactual effect of a truly non-selective mixed-attainment grouping. It is possible that the true effect of attainment grouping on children's verbal and psychological outcomes is masked. The differences in child outcomes between the comparator group (streamed or/and set) and the reference group (not streamed and not set) might be mitigated by within-class grouping. It may be likely that the differences in children's verbal skills and psychological symptoms by attainment grouping status are less pronounced since both groups may be exposed to within-class grouping during primary school.

8.5.2.2: Non-equivalence of streaming and setting

The operationalisation of attainment in this PhD rests on the assumption that setting, and streaming are equivalent. It was then difficult to classify the grouping arrangements within a school, where there may be several different types of operating at any one time (Hallam and Parsons, 2013b). Although setting and streaming are combined in the evidence reviews on attainment grouping (Education Endowment Foundation, 2018a), they constitute different forms of attainment grouping. Streaming is a rigid form of attainment grouping where students are placed into a class based on a measure of their overall ability and remain in that class for most curriculum subjects (Ireson and Hallam, 2001). In contrast, setting is a more flexible way of grouping based on their attainment in a curriculum subject (Ireson and Hallam, 2001). Recently, researchers from the UCL IoE suggested educators to avoid extrapolating setting across subjects, thus introducing elements of streaming (Francis et al., 2018). It is therefore possible that streaming may exert a stronger influence on children's learning and psychological outcomes than setting, which could be mitigated when the two are combined. Besides, it is also possible that the increase in the gap in verbal skills at school entry (high versus low scores) who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups is bigger among those who transitioned to streaming than those MCS participants who transitioned to setting.

8.5.2.3: Lack of information on attainment grouping before MCS4 (age 7)

Information on attainment grouping was provided by the MCS participants' teachers only at MCS4 (age 7) and MCS5 (age 11). It was not then possible to know whether MCS participants were grouped by attainment before MCS4 i.e at MCS3 (age 5). It is theoretically possible that MCS participants who were either classified as not being streamed/set at age 7 or were classified as not being streamed/set at both age 7 and age 11, might have been streamed/set between age 5 and age 7. Such a misclassification in the comparator group could have diluted the associations of interest. Besides, having data on attainment grouping from the beginning of primary school would have allowed a proper examination of the parallel trends' assumption for the statistical analyses of chapters 5 and 6.

8.5.2.4: Lack of information on the timing of attainment grouping transition

The lack of information on the timing of the transition from mixed to attainment grouping between ages 7 and 11 years is an important limitation. Some of MCS participants may have been taught in attainment groups as little as one week, while some others may have been taught in attainment groups up to four years. It is then possible that a dose response relationship exists between the length of attainment grouping and the subsequent change in verbal skills. This dose response relationship may be possibly masked when ignoring the timing of attainment grouping transition. The effect of being taught in an attainment group for a longer period may be mitigated by a shorter period of being taught in an attainment group. Besides, it is possible that the lack of verbal skills improvement between ages 7 and 11 years was more common among those MCS participants who entered primary school with lower levels of verbal skills, and they were taught for longer time in attainment groups than those MCS participants who were taught in attainment groups for shorter time between ages 7 and 11 years.

8.5.2.5: The transition to mixed attainment grouping between ages 7 and 11

The transition from mixed to attainment grouping between ages 7 and 11 years was not the only transition observed in the MCS. A considerable proportion of MCS participants (between 12-17%) transitioned from attainment grouping to mixed grouping, while 17-30% of MCS participants were taught in attainment groups at both ages 7 and 11 years. It would have been useful to explore the effect of transitioning from attainment grouping to mixed grouping on changes in verbal skills and psychological symptoms as compared to the counterfactual of continuously being taught in attainment groups at both ages 7 and 11. Nevertheless, provided the limited time and funding to complete this PhD topic rendered difficult such an endeavour. There is a paucity of studies exploring how a transition to mixed grouping might affect the cognitive and psychological development of children compared to continuous attainment grouping in primary school. Future studies could possibly inform educators and policy makers for the merits, if any, of a policy change towards mixed grouping teaching and for whom. It is possible that children who transition to mixed grouping experience a greater improvement in their verbal skills than children who are continuously taught in attainment groups. A transition to mixed grouping may be associated with being exposed to a diverse and rich curriculum, and collaborative learning which may lead to greater improvement in verbal skills and psychological symptoms for all children, and potentially benefit more children from disadvantaged socio-economic background or children who are less school ready (Tereshchenko et al., 2019). An arguable weakness of future MCS studies on the topic will be the lack of information on the length mixed grouping among those MCS participants who transitioned to mixed grouping between ages 7 and 11 years.

8.5.2.6: Unobserved time-varying confounding

This PhD used observational data and not experimental data in measuring the effect of attainment grouping in primary school. Despite the use of advanced statistical techniques in chapters 5 and 6, such as the fixed effects difference-in-difference modelling, aiming to reduce the room for omitted variable bias, time varying confounding remains an issue. A difference-in-difference design can consider shared changes between the group of students who changed to attainment grouping between 7 and 11 years and those who did not. However, it is unable to account for factors that could affect the reference and the comparator groups differently over time (e.g. school funding or school leadership).

8.5.3: Statistical power of three-way interactions

The three-way interactions between time (pre/post), attainment grouping, individual (baseline verbal skills) or family characteristics (family income) in chapters 5 and 6 could be underpowered, resulting from a combination of attrition rates from the start of the MCS (~28%) (Hansen et al., 2014) and the non-response to the MCS4 (~30%)and MCS5 (~23%) teacher surveys (Mostafa and Rosenberg, 2013, Gallop et al., 2015). Therefore, these interactions should be perceived as exploratory. Caution should be exercised when interpreting the statistically significant increase in verbal skills differences from age 7 to age 11 in England and Wales, among MCS participants with different levels of verbal skills at age 5 and who changed to attainment grouping between ages 7 and 11 years (Heo and Leon, 2010). This result was highly imprecise. The three-way interaction term (attainment grouping transition*initial verbal skills*time) ranged from 0.02 to 0.19 unit increase in verbal differences for each unit increase in baseline verbal skills for those who changed to attainment grouping (see Table A2.14 in Appendix 2), despite a high degree of accuracy in measuring verbal skills through the British Ability Scales and a large enough sample.

8.5.4: Differences at the extremes of the outcome distributions

This PhD focused on attainment grouping effects on mean verbal skills and psychological symptoms over time. Provided the limited time and funding for the completion of the current PhD topic, it was not possible to explore possible effects of attainment grouping at the extremes of these developmental outcomes. Currently, there is a paucity of such studies in the literature of attainment grouping. So, future studies deviating from mean "population" effects could fill this gap. It may be possible that attainment grouping in primary school may be related to an increased likelihood of clinically relevant verbal and psychological difficulties. The MCS provides an ideal opportunity for such research given the availability of validated and clinically relevant cut-offs for verbal skills and psychological symptoms (Tomblin et al., 1997, Goodman, 1999).

8.5.5: Data limitations

8.5.5.1: Verbal skills

It should be noted that the BAS subtests at MCS 3, MCS 4 and MCS 5 did not measure consistently the same verbal cognitive processes. At MCS 3 the BAS Naming Vocabulary subscale measured children's expressive language and vocabulary. At MCS 4 the BAS Word Reading subscale measured children's reading ability, and at MCS 5 the BAS Verbal Similarities subtest measured children's verbal reasoning and verbal knowledge. Nevertheless, it is plausible that BAS subscales are measuring different aspects of what are perhaps overlapping skills of the same underlying construct, verbal skills. Prior studies using the MCS data (Zilanawala, 2016, Zilanawala et al., 2017, Bradbury et al., 2015) have used the different BAS subscales to understand changes in this underlying construct over time. The BAS subscale scores used in this PhD were standardized to mean 50 and standard deviation 10 and are adjusted for both item difficulty and age. This was done to get a satisfactory general measure of children's verbal cognitive development from MCS 3 to MCS 5. However,

8.5.5.2: Family income

Family income is arguably the best single indicator of material living standards. However, there are several limitations to the MCS family income measure that ought to be acknowledged. The family income measure available in the MCS used only a single banded question of different size for single- and two-parent families, thus it provides a rough approximation, but not the actual amount of respondents' income. Previous studies have shown that using single measures of income could yield less accurate estimates of the actual income; that is respondents tend to underreport or select an income band below which their income actually falls (Hansen and Kneale, 2013). Analysis has shown that the MCS family income measure did not match well with data from the Households Below Average Income statistics (Hansen et al., 2014), a survey with several income questions, thus questioning the accuracy of the MCS derived family income variable. However, the MCS income measure has been found to fare well in terms of the relative differences between groups and countries (Hansen et al., 2014).

A further reason for caution regarding the accuracy of the MCS family income measure is the omission of state benefits such as Housing Benefit and Council Tax Benefit in the estimates of the total family net income (Hansen et al., 2014, Fitzsimons et al., 2017a).

8.5.5.3: Lack of data on numeracy skills

In the MCS, longitudinal information during primary school was only available for verbal skills, but not for numeracy skills. This PhD found that a transition to attainment grouping (streaming/setting for literacy) between ages 7 and 11 years was not associated with an overall change in verbal skills. However, it could be that a transition to attainment grouping (streaming setting for numeracy) is beneficial or detrimental to children's numerical skills during primary school. Further studies, which take progress into numerical skills into account, will need to be undertaken.

8.5.5.4: Lack of data on pedagogy

This PhD was not able to provide an explanation/answer to the finding of chapter 5 which highlighted a widening gap in verbal skills differences from ages 7 to 11 years between children with high and low initial levels of verbal skills. This could be possible if the MCS had data on pedagogy, teaching expectations or teaching styles that could particularly account for the lack of verbal skills improvement from age 7 to age 11 for those MCS participants who started primary school in England and Wales with a lower level of verbal skills and were subsequently grouped by attainment between ages 7 and 11 years.

8.5.5.5: Lack of data on the number of and the time spent in attainment groups

There is lack of information in the MCS regarding the quantity of attainment grouping during primary school (in how many groups MCS participants were taught and for how long).

In the MCS4, teachers were asked whether MCS participants were streamed or set for core curriculum subjects (English or Maths), while during MCS5 were asked about streaming/setting in English, Maths and Science. It is likely that some MCS participants were in attainment groups for core curriculum subjects or Science, but they were also taught in mixed attainment groups for Topic subjects (Finally, the MCS lacks information on the time spent in streams/sets. Thus, it was not possible to explore dose-response relationships between the amount of time spent in streams/sets for certain or all curriculum subjects during primary school and MCS participants cognitive and psychological outcomes.

8.6: Future research

There are different areas that, by building on and expanding this PhD, could be explored in future research. Some of them, mentioned in what follows, might address some of the limitations of this research.

8.6.1: Standardised achievement tests in the MCS

This PhD showed that there was not a significant overall effect of the transition to attainment grouping between ages 7 and 11 years on the development of children's verbal skills. There was, however, evidence of a widening in the gap in verbal skills at school entry (high versus low scores) for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups. The gap in verbal skills at school entry (high versus low scores) widened by approximately 4.35 BAS points for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed attainment groups at ages 7 and 11 years, after statistical adjustment for their transition to top, middle and bottom streams or/and sets for literacy and maths. Provided that BAS has a standard deviation of 10, there was approximately a 40% or 0.40 of a standard deviation increase in the verbal skills gap (top vs. bottom quintile at age 5) by the end of primary school linked to the transition to attainment grouping between ages 7 and 11 years. Available evidence shows that school interventions during primary/elementary school have an average effect size ranging from 0.20 to ~0.35 standard deviations (Lipsey et al., 2012). Considering the above, even an approximately 0.40 standard deviation increase is probably of educational significance since attainment grouping is amenable to policy influence at a minimal cost (Duncan and Magnuson, 2007, Levin and Belfield, 2015, Education Endowment Foundation, 2018a).

KS test scores from the NPD are available for MCS participants at age 7 and age 11 under secure arrangements via the UK Data Service. Future research should exploit this opportunity to replicate the findings of chapter five. Fixed effects difference-in-difference modelling could be applied to explore whether a transition to attainment grouping between ages 7 and 11 years is associated with progress in English and Maths from KS1 (age seven) to KS2(age eleven). Future research could also go a step further and investigate whether attainment grouping in primary school has an effect on GSCE assessment at age 16, provided that GSCE data from the NPD have recently been linked to the MCS (Centre for Longitudinal Studies and Department for Education., 2019).

8.6.2: Estimating differences at the extremes of outcome distribution

This PhD did not find evidence for an overall effect of attainment grouping in primary school on children's and adolescents' average/mean verbal skills, externalising or internalising psychological symptoms. Future research could apply multilevel logistic regression to examine associations between attainment grouping in primary school and probable child psychiatric caseness or probable child language problems. Multilevel quantile regression modelling could also be applied to examine whether the magnitude of inequalities of differences (e.g socioeconomic or /cognitive) differs across quantiles of the of the externalising or internalising psychological symptoms distribution.

8.6.3: Non-cognitive/educational outcomes

Given the scarce evidence on the effect of attainment grouping in primary school on non-cognitive/educational outcomes, future research could use a variety of outcome measures available in the MCS, such as self-esteem, academic self-concept and attitudes towards school to investigate the wider consequences of attainment grouping in primary school.

8.6.4: Effects of attainment grouping by other socio-economic measures

Given the limitations of the MCS family income measure, future research using the MCS can benefit by using other measures of socio-economic circumstances to explore potential differential effects of attainment grouping on children's outcomes. Markers of financial hardship in the family may be better suited in capturing the varied challenges in families' lives (Mortimer et al., 2014, Beck et al., 2014). In the MCS financial hardship in the family can be measured via reporting being behind with bills, experience of current unemployment or how well the family is managing financially.

8.6.5: Effects of attainment grouping by ethnicity, SEN or EAL

The greater part of attainment grouping literature pays particular attention to overall/mean effects of attainment grouping or its role on increasing educational differences (Education Endowment Foundation, 2018a). To develop a full picture of the possible moderating effects of attainment grouping on children's cognitive and psychological outcomes, additional studies considering several child characteristics are needed. In England and Wales, approximately a third of primary school pupils are from ethnic minority groups, a fifth of primary school pupils are exposed to a language other than English at home (EAL) and approximately a quarter of primary school pupils have SEN (Office for National Statistics., 2019, Department for Education., 2019b). Evidence also shows differences in cognitive and psychological outcomes of children from different ethnic groups or between children with EAL or SEN (Whiteside et al., 2017, Emerson and Hatton, 2007, Zilanawala, 2016, Zilanawala et al., 2015, Strand, 2014, Portes and MacLeod, 1996, Leventhal et al., 2006, Goodman et al., 2008, Belfi et al., 2014). The MCS includes high quality data on such characteristics, so It would be interesting to explore in the future whether attainment grouping in primary school amplifies or mitigates such differences.

8.6.6: Time-varying confounding

Whilst future research on attainment grouping should continue to consider the potential effects for cognitive/academic and non-cognitive outcomes, it should also consider the possible methodological issues associated with estimating the effect of attainment grouping. Issues concerning the effective adjustment for time-invariant and time-varying confounding have been addressed within a very limited number of studies in the attainment grouping literature. An attempt was made to control for time-invariant characteristics as reported by MCS participant's teacher, however it should be noted that some of the controls might be imperfect proxies for the true confounding variable. For instance, the number of classes within a child's school year might be an imperfect proxy for the child's school size. Time-varying confounding at the school-level was also an issue in this PhD. That is some unobserved school characteristics e.g. change in school funding over time or teacher turnover may still explain the increase in children's initial verbal skills differences related to a transition to attainment grouping between MCS 4 and MCS 5. Future studies using the MCS will be benefited by linking rich timely data from the DfE such as the distribution of financial resources to UK primary schools between 2008 and 2012 or teacher retention rates in UK primary schools between 2008 and 2012. By complementing the statistical analyses of this PhD and accounting for such data, future research will be able to provide more robust and compelling evidence on the overall and heterogeneous effects of attainment grouping on children's cognitive/educational, psychological and other non-cognitive outcomes.

8.7 Concluding remarks

This PhD set out to investigate associations between attainment grouping in UK primary schools and the development of MCS participants' verbal skills and psychological symptoms in childhood and adolescence. There was no evidence for an overall effect of attainment grouping transition between ages 7 and 11 years on children's verbal skills from age 7 to age 11 in English and Welsh primary schools. Neither was evidence found of an overall effect of attainment grouping transition on MCS participants' externalising and internalising psychological symptoms from age 7 to age 11. No evidence was found on an association between attainment grouping (age seven) and MCS participants' externalising and internalising psychological symptoms from age 7 to age 14. Attainment grouping in primary school also seemed to play no or little role in socioeconomic and educational related inequalities in children's and adolescents 'psychological symptoms in the MCS.

In contrast, the gap in verbal skills at school entry (high versus low scores) widened for children who transitioned from mixed to attainment groups between age 7 and 11 compared to those who remained in mixed groups.

What primary schools do in the UK may exacerbate children's early home and early education experiences, thus contributing to an accumulation of disadvantage among low-attaining children and an increase in verbal cognitive differences over the sensitive period of childhood. Future studies using the MCS and other UK surveys of primary school aged children could benefit from linking administrative and school census data to adjust for time varying characteristics between schools and explore more rigorously the existence of a causal relationship between attainment grouping and an increase in inequalities in children's cognitive and academic attainment. If causal links are established, it is of utmost importance to understand the exact pathways underpinning these inequalities. In the meantime, school policy makers and teachers should continue doing what works best for their school by ensuring that placement to attainment groups reflects pupils' attainment and not their social background or their individual characteristics, by ensuring equal access for all students to a rich curriculum and by applying high expectations to all attainment groups.

Bibliography

ACHENBACH, T. M. & EDELBROCK, C. S. 1978. Classification of Child Psychopathology - Review and Analysis of Empirical Efforts. *Psychological Bulletin,* 85**,** 1275-1301.

AFENDULIS, C. C., HE, Y. L., ZASLAVSKY, A. M. & CHERNEW, M. E. 2011. The Impact of Medicare Part D on Hospitalization Rates. *Health Services Research,* 46**,** 1022-1038.

AHN, J. V., SERA, F., CUMMINS, S. & FLOURI, E. 2018. Associations between objectively measured physical activity and later mental health outcomes in children: findings from the UK Millennium Cohort Study. *Journal of Epidemiology and Community Health,* 72**,** 94-100.

ALLEN, R., BURGESS, S. & MAYO, J. 2012. The teacher labour market, teacher turnover and disadvantaged schools: new evidence for England. Bristol: The Centre for Market and Public Organisation, University of Bristol.

ALLISON, P., D. 2009. *Fixed Effects Regression Models*, SAGE publications.

ANASTASI, A. 1982. *Psychological testing.,* New York, Macmillan.

ANDREWS, J., HUTCHINSON, J. & JOHNES, R. 2016. Grammar schools and social mobility. London: Education Policy Institute.

ANGRIST, J. D. & PISCHKE, J.-S. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, NJ, Princeton University Press.

BAINES, E., BLATCHFORD, P. & KUTNICK, P. 2003. Changes in grouping practices over primaryand secondary school. *International Journal of Educational Research,* 39**,** 9-34.

BALL, S. 1981. *Beachside comprehensive: A case-study of secondary schooling.,* Cambridge, Cambridge University Press.

BALL, S. J., BOWE, R. & GEWIRTZ, S. 1994. Competitive Schooling: Values, Ethics, and Cultural Engineering. *Journal of Curriculum and Supervision,* 9**,** 350-367.

BANGAY, C. & BLUM, N. 2010. Education responses to climate change and quality: Two parts of the same agenda? *International Journal of Educational Development,* 30**,** 359-368.

BARBEY, A. K. 2018. Network Neuroscience Theory of Human Intelligence. *Trends in Cognitive Sciences,* 22**,** 8-20.

BARKER, D. J. 1990. The fetal and infant origins of adult disease. *BMJ,* 301**,** 1111.

BARTLEY, M. 2017. *Health Inequality: An Introduction to Concepts, Theories and Methods (Second Edition)*, Polity

BECK, A. F., HUANG, B., SIMMONS, J. M., MONCRIEF, T., SAUERS, H. S., CHEN, C., RYAN, P. H., NEWMAN, N. C. & KAHN, R. S. 2014. Role of financial and social hardships in asthma racial disparities. *Pediatrics,* 133**,** 431-9.

BELFI, B., GOOS, M., PINXTEN, M., VERHAEGHE, J. P., GIELEN, S., DE FRAINE, B. & VAN DAMME, J. 2014. Inequality in language achievement growth? An investigation into the impact of pupil socio-ethnic background and school socio-ethnic composition. *British Educational Research Journal,* 40**,** 820-846.

BEN-SHLOMO, Y. & KUH, D. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol,* 31**,** 285-93.

BERCOW, J. 2008. The Bercow report- A Review of Services for Children and Young People (0-19) with Speech, Language and Communication Needs.

BERENDS, M. 1995. Educational Stratification and Students' Social Bonding to School. *British Journal of Sociology of Education* 16**,** 327-351.

BLACKSTONE, T. 1967. Plowden Report. *British Journal of Sociology,* 18**,** 291-302.

BLANDEN, J., GREAVES, E., GREGG, P., MACMILLAN, L. & SIBIETA, L. 2015. Understanding the improved performance of disadvantaged pupils in London. . *CASE SPCC Working Paper No.21.* . London: London School of Economics.

BLATCHFORD, P., BASSETT, P. & BROWN, P. 2011. Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction,* 21**,** 715-730.

BLATCHFORD, P., BASSETT, P., GOLDSTEIN, H. & MARTIN, C. 2003. Are class size differences related to pupils' educational progress and classroom processes? Findings from the institute of education class size study of children aged 5-7 years. *British Educational Research Journal,* 29**,** 709-730.

BOALER, J. 1997a. Setting, Social Class and Survival of the Quickest. *British Educational Research Journal,* 23**,** 575-595.

BOALER, J. 1997b. When even the winners are losers: evaluating the experience s of 'top set' students. *Journal of Curriculum Studies,* 29**,** 165-182.

BOALER, J., WILIAM, D. & BROWN, M. 2000. Students' Experiences of Ability Grouping—disaffection, polarisation and the construction of failure. *British Educational Research Journal,* 26**,** 631-648.

BORRELL-PORTA, M., COOPER, K., COSTA FONT, J., ORSINI, C. & PLATT, L. 2017. Children's Wellbeing And Development Outcomes For Ages 5, 7 and 11, and their Predictors (Final Report)

London School of Economics.

BOTTING, N., CONTI-RAMSDEN, G. 2000. Child language and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy,* 16**,** 105-120.

BOWES, L., ARSENEAULT, L., MAUGHAN, B., TAYLOR, A., CASPI, A. & MOFFITT, T. E. 2009. School, Neighborhood, and Family Factors Are Associated With Children's Bullying Involvement: A Nationally Representative Longitudinal Study. *Journal of the American Academy of Child and Adolescent Psychiatry,* 48**,** 545-553.

BRADBURY, A. & ROBERTS-HOLMES, G. 2017. Grouping in Early Years and Key Stage 1. London: National Education Union & UCL Institute of Education.

BRADBURY, B., CORAK, M., WALDFOGEL, J. & WASHBROOK, E. 2015. *Too Many Children Left Behind: The U.S. Achievement Gap in Comparative Perspective*, Russell Sage Foundation

BRADBURY, B., WALDFOGEL, J. & WASHBROOK, E. 2019. Income-Related Gaps in Early Child Cognitive Development: Why Are They Larger in the United States Than in the United Kingdom, Australia, and Canada? *Demography* 56**,** 367-390.

BRANAS, C. C., CHENEY, R. A., MACDONALD, J. M., TAM, V. W., JACKSON, T. D. & TEN HAVE, T. R. 2011. A Difference-in-Differences Analysis of Health, Safety, and Greening Vacant Urban Space. *American Journal of Epidemiology,* 174**,** 1296-1306.

BRANDT, R. & BRONFENBRENNER, U. 1979. On Families and Schools - Conversation with Bronfenbrenner,Urie. *Educational Leadership,* 36**,** 459-463.

BRONFENBRENNER, U. 1977. Toward an Experimental Ecology of Human-Development. *American Psychologist,* 32**,** 513-531.

BRONFENBRENNER, U. 1979. Contexts of Child-Rearing - Problems and Prospects. *American Psychologist,* 34**,** 844-850.

BRONFENBRENNER, U. & EVANS, G. W. 2000. Developmental science in the 21st century: Emerging questions, theoretical models, research designs and empirical findings. *Social Development,* 9**,** 115-125.

BRONFENBRENNER, U., MOEN, P. & GARBARINO, J. 1984. Child, Family, and Community. *Review of Child Development Research,* 7**,** 283-328.

BROOKOVER, W. B., SCHWEITZER, J. H., SCHNEIDER, J. M., BEADY, P. K., FLOOD, P. K. & WISENBAKER, J. M. 1978. Elementary school social climate and school achievement *American Educational Research Journal,* 15**,** 301–318.

BROWN, P., GREEN, A. & LAUDER, H. 2001. *High Skills: Globalisation, Competitiveness and Skills Formation,* Oxford, Oxford University Press.

BRUNELLO, G. & CHECCHI, D. 2006. Does school tracking affect equality of opportunity? New International Evidence. . *IZA Discussion Paper 2348.* Institute for the Study of Labor (IZA).

BURGESS, S., GREGG, P., , HALL, E., MEADOWS, S., PROUD, S., , PROPPER, C., & WASHBROOK, L.
2006. Up to age 7: family background and child development up to age 7 in the Avon L ongitudinal Survey of Parents and Children (ALSPAC). Bristol: University of Bristol.

CAMPBELL, T. 2015. Stereotyped at Seven? Biases in Teacher Judgement of Pupils' Ability and Attainment. *Journal of Social Policy,* 44**,** 517-545.

CAMPBELL, T., GAMBARO, L. & STEWART, K. 2019. Inequalities in the experience of early education in England: access, peer groups and transitions. *CASEpapers (214). Centre for Analysis of Social Exclusion.* London:
School of Economics and Political Science.

CARD, D. & KRUEGER, A. B. 1994. Minimum-Wages and Employment - a Case-Study of the Fast-Food Industry in New-Jersey and Pennsylvania. *American Economic Review,* 84**,** 772-793.

CARLSSON, M., DAHL, G. B., ÖCKERT, B. & ROOTH, D.-O. 2015. The Effect of Schooling on Cognitive Skills *Review of Economics and Statistics,* 97**,** 533-547.

CARNEIRO, P. & HECKMAN, J. J. 2003. Human Capital Policy. *In:* HECKMAN, J. J., KRUEGER, A. B. & FRIEDMAN. B.M. (eds.) *In Inequality in America: What Role for Human Capital Policies ?* Cambridge, MA: MIT Press.

CARPENTER, J. R. & KENWARD, G. M. 2013. *Multiple Imputation and its Application*, John Wiley & Sons.

CENTRE FOR LONGITUDINAL STUDIES AND DEPARTMENT FOR EDUCATION. 2019. Millennium Cohort Study: Linked Education Administrative Datasets (National Pupil Database), England: Secure Access.UK Data Service. . UK Data Service.

CHANG, C. K., HAYES, R. D., PERERA, G., BROADBENT, M. T. M., FERNANDES, A. C., LEE, W. E., HOTOPF, M. & STEWART, R. 2011. Life Expectancy at Birth for People with Serious Mental Illness and Other Major Disorders from a Secondary Mental Health Care Case Register in London. *Plos One,* 6.

CHETTY, R., LOONEY, A. & KROFT, K. 2009. Salience and Taxation: Theory and Evidence. *American Economic Review,* 99**,** 1145-1177.

CHZHEN, Y., REES, G., GROMADA, A., CUESTA, J. & Z., B. 2018. An Unfair Start-Inequality in Children's Education in Rich Countries. *In:* DROHAN, M. (ed.) *Innocenti Report Card 15.* Florence: UNICEF, Office of Research-Innocenti.

CILLESSEN, A. H. N. & BORCH, C. 2006. Developmental trajectories of adolescent popularity: A growth curve modelling analysis. *Journal of Adolescence,* 29**,** 935-959.

CLARK, C., SMUK, M., LAIN, D., STANSFELD, S. A., CARR, E., HEAD, J. & VICKERSTAFF, S. 2017. Impact of childhood and adulthood psychological health on labour force participation and exit in later life. *Psychol Med,* 47**,** 1597-1608.

CLEGG, J., HOLLIS, C., MAWHOOD, L. & RUTTER, M. 2005. Developmental language disorders a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry,* 46**,** 128-149.

CLIFTON, J. & COOK, W. 2012. A long division: Closing the attainment gap in England's secondary schools. London: IPPR.

COLMAN, I., MURRAY, J., ABBOTT, R. A., MAUGHAN, B., KUH, D., CROUDACE, T. J. & JONES, P. B. 2009. Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. *Bmj-British Medical Journal,* 338.

COLMAN, I., WADSWORTH, M. E. J., CROUDACE, T. J. & JONES, P. B. 2007. Forty-year psychiatric outcomes following assessment for internalizing disorder in adolescence. *American Journal of Psychiatry,* 164**,** 126-133.

CONNELY, R. 2013. Millennium Cohort Study Data Note 2013/1. *Interpreting Test Scores.* London: Centre for Longitudinal Studies.
CONTI-RAMSDEN, G., DURKIN, K., SIMKIN, Z. & KNOX, E. 2009. Specific language impairment and school outcomes. I: Identifying and explaining variability at the end of compulsory education. *International Journal of Language & Communication Disorders,* 44**,** 15-35.
CONTI-RAMSDEN, G., DURKIN, K., TOSEEB, U., BOTTING, N. & PICKLES, A. 2018. Education and employment outcomes of young adults with a history of developmental language disorder. *Int J Lang Commun Disord,* 53**,** 237-255.
COOPE, C. M. 2013. *Maternal psychological distress and early-childhood externalising behaviour : a longitudinal analysis of the UK Millennium Cohort Study* P.h.D, University College London.
COOPER, K. 2017. *Poverty and Parenting in the UK.* Doctor of Philosophy, London School of Economics and Political Science.
COOPER, K. & STEWART, K. 2013a. Does money affect children's outcomes. London: Joseph Rowntree Foundation.
COOPER, K. & STEWART, K. 2013b. Does money affect children's outcomes? London: Joseph Rowntree Foundation.
COOPER, K. & STEWART, K. 2017. Does Money Affect Children’s Outcomes? An update. *CASEpaper 203.* Centre for Analysis of Social Exclusion, London School of Economics.
CRAIG, P., COOPER, C., GUNNELL, D., HAW, S., LAWSON, K., MACINTYRE, S., OGILVIE, D., PETTICREW, M., REEVES, B., SUTTON, M. & THOMSON, S. 2012. sing natural experiments to evaluate population health interventions: new Medical Research Council guidance. *Journal of Epidemiology and Community Health,* 66**,** 1182–1186.
CRAWFORD, C., MACMILLAN, L. & VIGNOLES, A. 2017. When and why do initially high-achieving poor children fall behind? *Oxford Review of Education,* 43**,** 88-108.
CREESE, H.-M. 2019. *Psychosocial stressors in the home environment, obesity and socioemotional difficulties across the first decade of life: Findings from the Millennium Cohort Study* P.h.D, University College London.
CROZIER, G., REAY, D. & JAMES, D. 2011. Making it work for them: White middle class families and working class schools. *International Studies in Sociology of Education,* 21**,** 199-216.
CURRAN, P. J., OBEIDAT, K. & LOSARDO, D. 2010. Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognition and Development,* 11**,** 121-136.
CURRIE, J. 2001. Early childhood education programs. *Journal of Economic Perspectives,* 15**,** 213-238.
DAVIES, G., LAM, M., HARRIS, S. E., TRAMPUSH, J. W., LUCIANO, M., HILL, W. D., HAGENAARS, S. P., RITCHIE, S. J., MARIONI, R. E., FAWNS-RITCHIE, C., LIEWALD, D. C. M., OKELY, J. A., AHOLA-OLLI, A. V., BARNES, C. L. K., BERTRAM, L., BIS, J. C., BURDICK, K. E., CHRISTOFOROU, A., DEROSSE, P., DJUROVIC, S., ESPESETH, T., GIAKOUMAKI, S., GIDDALURU, S., GUSTAVSON, D. E., HAYWARD, C., HOFER, E., IKRAM, M. A., KARLSSON, R., KNOWLES, E., LAHTI, J., LEBER, M., LI, S., MATHER, K. A., MELLE, I., MORRIS, D., OLDMEADOW, C., PALVIAINEN, T., PAYTON, A., PAZOKI, R., PETROVIC, K., REYNOLDS, C. A., SARGURUPREMRAJ, M., SCHOLZ, M., SMITH, J. A., SMITH, A. V., TERZIKHAN, N., THALAMUTHU, A., TROMPET, S., LEE, S. J. D., WARE, E. B., WINDHAM, B. G., WRIGHT, M. J., YANG, J., YU, J., AMES, D., AMIN, N., AMOUYEL, P., ANDREASSEN, O. A., ARMSTRONG, N. J., ASSAREH, A. A., ATTIA, J. R., ATTIX, D., AVRAMOPOULOS, D., BENNETT, D. A., BOEHMER, A. C., BOYLE, P. A., BRODATY, H., CAMPBELL, H., CANNON, T. D., CIRULLI, E. T., CONGDON, E., CONLEY, E. D., CORLEY, J., COX, S. R., DALE, A. M., DEHGHAN, A., DICK, D., DICKINSON, D., ERIKSSON, J. G., EVANGELOU, E., FAUL, J. D., FORD, I., FREIMER, N. A., GAO, H., GIEGLING, I., GILLESPIE, N. A., GORDON, S. D., GOTTESMAN, R. F., GRISWOLD, M. E., GUDNASON, V., HARRIS, T. B., HARTMANN, A. M., HATZIMANOLIS, A., HEISS, G., HOLLIDAY, E. G., JOSHI, P. K., KAHONEN, M., KARDIA, S. L. R., KARLSSON, I., KLEINEIDAM, L., et al. 2018. Study of 300,486 individuals

identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications,* 9.
DEARDEN, L., SIBIETA, L. & SYLVA, K. 2011. The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study. *Longitudinal and Life Course Studies* 2**,** 19-40.
DEARY, I. J., JOHNSON, W. & HOULIHAN, L. M. 2009. Genetic foundations of human intelligence. *Human Genetics,* 126**,** 215-232.
DEL BONO, E., FRANCESCONI, M., KELLY, Y. & SACKER, A. 2016. Early Maternal Time Investment and Early Child Outcomes. *Economic Journal,* 126**,** F96-F135.
DEPARTMENT FOR EDUCATION 2017. Analysis of teacher supply, retention and mobility.
DEPARTMENT FOR EDUCATION AND EMPLOYMENT. 1997. Excellence in schools. London: DfEE.
DEPARTMENT FOR EDUCATION. 2010. The importance of teaching: The Schools White Paper 2010. London: DfE.
DEPARTMENT FOR EDUCATION. 2017. GSCE results 2017.
DEPARTMENT FOR EDUCATION. 2019a. Schools, pupils and their characteristics: June 2019.
DEPARTMENT FOR EDUCATION. 2019b. Special educational needs: an analysis and summary of data sources
DICKERSON, A. & POPLI, G. K. 2016. Persistent poverty and children's cognitive development: evidence from the UK Millennium Cohort Study *Journal of the Royal Statistical Society,* 179**,** 535-558.
DIMICK, J. B. & RYAN, A. M. 2014. Methods for Evaluating Changes in Health Care Policy The Difference-in-Differences Approach. *Jama-Journal of the American Medical Association,* 312**,** 2401-2402.
DUFLO, E., DUPAS, P. & KREMER, M. 2011. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review,* 101**,** 1739-1774.
DUNCAN, G. J. & MAGNUSON, K. 2007. Penny wise and effect size foolish. *Child Development Perspectives,* 1**,** 46-51.
DUNCAN, G. J., MAGNUSON, K. & VOTRUBA-DRZAL, E. 2014. Boosting Family Income to Promote Child Development. *Future of Children,* 24**,** 99-120.
DWECK, C. S. 2007. *Mindset: The New Psychology of Success*, Ballantine Books.
EDDINGS, W. & MARCHENKO, Y. 2012. Diagnostics for multiple imputation in Stata. *The Stata Journal,* 12**,** 353-367.
EDUCATION ENDOWMENT FOUNDATION 2018a. Setting or Streaming.
EDUCATION ENDOWMENT FOUNDATION 2018b. Within-class attainment grouping.
EGAN, M., DALY, M. & DELANEY, L. 2016. Adolescent psychological distress, unemployment, and the Great Recession: Evidence from the National Longitudinal Study of Youth 1997. *Soc Sci Med,* 156**,** 98-105.
ELDER, G. H. 1998. The life course as developmental theory. *Child Development,* 69**,** 1-12.
ELGAR, F. J., CRAIG, W. & TRITES, S. J. 2013. Family dinners, communication, and mental health in Canadian adolescents. *J Adolesc Health,* 52**,** 433-8.
ELLIOT, C. D., SMITH, P. & MCCULLOGH, K. 1997. British Ability Scales Second Edition (BAS II). Technical Manual. London.
EMERSON , E. & HATTON, C. 2007. The Mental Health of Children and Adolescents with Learning Disabilities in Britain. Institute for Health Research.
ENDERS, C. K. 2010. *Applied Missing Data Analysis,* New York, The Guilford Press.
ESCH, P., BOCQUET, V., PULL, C., COUFFIGNAL, S., LEHNERT, T., GRAAS, M., FOND-HARMANT, L. & ANSSEAU, M. 2014. The downward spiral of mental disorders and educational attainment: a systematic review on early school leaving. *Bmc Psychiatry,* 14.
FAHEY, S. J. 2012. Curriculum change and climate change: Inside outside pressures in higher education. *Journal of Curriculum Studies,* 44**,** 703-722.

FANCOURT, D. & STEPTOE, A. 2019. Effects of creativity on social and behavioral adjustment in 7- to 11-year-old children. *Ann N Y Acad Sci,* 1438**,** 30-39.
FEINSTEIN, L. & BYNNER, J. 2004. The importance of cognitive development in middle childhood for adulthood socioeconomic status, mental health, and problem behavior. *Child Dev,* 75**,** 1329-1339.
FIESE, B. H. & SCHWARTZ, M. 2008. Reclaiming the family table: Mealtimes and child health and wellbeing. *Social Policy Report,* 22.
FINN, J. D., PANNOZZO, G. M. & ACHILLES, C. M. 2003. The "why's" of class size: Student behavior in small classes. *Review of Educational Research,* 73**,** 321-368.
FITZSIMONS, E., AGALIOTI-SGOMPOU, V., CALDERWOOD, L., GILBERT, E., HASELDEN, L., JOHNSON, J., SMITH, K. & TEAM, M. C. 2017a. Millennium Cohort Study; Sixth Survey 2015-2016-User Guide (First Edition). London, UCL Institute for Education: Centre for Longitudinal Studies.
FITZSIMONS, E., GOODMAN, A., KELLY, E. & SMITH, J. 2017b. Correlates of Mental Illness and Wellbeing in Children: Are They the Same? Results From the UK Millennium Cohort Study. *Social Science and Medicine,* 175**,** 43-51.
FLOURI, E., MAVROVELI, S. & TZAVIDIS, N. 2012. Cognitive ability, neighborhood deprivation, and young children’s emotional and behavioral problems. *Social Psychiatry and Psychiatric Epidemiology,* 47**,** 985-992.
FLOURI, E. & MIDOUHAS, E. 2016. School composition, family poverty and child behaviour. *Social Psychiatry and Psychiatric Epidemiology,* 51**,** 817-826.
FLOURI, E. & MIDOUHAS, E. 2017. Environmental Adversity and Children's Early Trajectories of Problem Behavior: The Role of Harsh Parental Discipline. *Journal of Family Psychology,* 31**,** 234-243.
FLOURI, E., MIDOUHAS, E. & FRANCESCONI, M. 2019a. Neighbourhood deprivation and child behaviour across childhood and adolescence. *Longitudinal and Life Course Studies*.
FLOURI, E., MIDOUHAS, E. & JOSHI, H. 2014. Family poverty and trajectories of children's emotional and behavioural problems: the moderating roles of self-regulation and verbal cognitive ability. *J Abnorm Child Psychol,* 42**,** 1043-56.
FLOURI, E., MIDOUHAS, E. & JOSHI, H. 2015a. Family and neighbourhood risk and children's problem behaviour: The moderating role of intelligence. *Intelligence,* 53**,** 33-42.
FLOURI, E., MIDOUHAS, E., JOSHI, H. & SULLIVAN, A. 2015b. Neighbourhood social fragmentation and the mental health of children in poverty. *Health & Place,* 31**,** 138-145.
FLOURI, E., MIDOUHAS, E., JOSHI, H. & TZAVIDIS, N. 2015c. Emotional and behavioural resilience to multiple risk exposure in early life: the role of parenting. *Eur Child Adolesc Psychiatry,* 24**,** 745-55.
FLOURI, E., MIDOUHAS, E. & RUDDY, A. 2016. Socio-economic status and family structure differences in early trajectories of child adjustment: Individual and neighbourhood effects. *Health & Place,* 37**,** 8-15.
FLOURI, E., MIDOUHAS, E. & TZATZAKI, K. 2015d. Neighbourhood and own social housing and early problem behaviour trajectories. *Social Psychiatry and Psychiatric Epidemiology,* 50**,** 203-213.
FLOURI, E., PAPACHRISTOU, E. & MIDOUHAS, E. 2019b. The role of neighbourhood greenspace in children's spatial working memory. *British Journal of Educational Psychology,* 89**,** 359-373.
FLOURI, E., RUDDY, A. & MIDOUHAS, E. 2017. Maternal depression and trajectories of child internalizing and externalizing problems: the roles of child decision making and working memory. *Psychological Medicine,* 47**,** 1138-1148.
FLOURI, E. & SARMADI, Z. 2016. Prosocial Behavior and Childhood Trajectories of Internalizing and Externalizing Problems: The Role of Neighborhood and School Contexts. *Developmental Psychology,* 52**,** 253-258.

FLOURI, E., SARMADI, Z. & FRANCESCONI, M. 2019c. Paternal Psychological Distress and Child Problem Behavior From Early Childhood to Middle Adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry,* 58**,** 453-458.

FRANCIS, B., ARCHER, L., HODGEN, J., PEPPER, D., TAYLOR, B. & TRAVERS, M. C. 2017a. Exploring the relative lack of impact of research on 'ability grouping' in England: a discourse analytic account. *Cambridge Journal of Education,* 47**,** 1-17.

FRANCIS, B., ARCHER, L., MAZENOD, A., CRAIG, N., TAYLOR, B., TERESHCHENKO, A., HODGEN, J. & CONNOLLY, P. 2019. Teacher 'quality' and attainment grouping: the role of within-school teacher deployment in social and educational inequality. *Teaching and Teacher Education,* 77**,** 183-192.

FRANCIS, B., CONNOLLY P, ARCHER L, HODGEN J, MAZENOD A, PEPPER D, SLOAN S, TAYLOR B, TERESHCHENKO A & M, T. 2017b. Attainment Grouping as self-fulfilling prophesy? A mixed methods exploration of self confidence and set level among Year 7 students. *International Journal of Educational Research,* 86**,** 96-108.

FRANCIS, B., TAYLOR, B., HODGEN, J., TERESHCHENKO, A. & ARCHER, L. 2018. Dos and Don'ts of attainment grouping. *In:* EDUCATION, U. I. O. (ed.). London.

GALAMBOS, N. L., BARKER, E. T. & ALMEIDA, D. M. 2003. Parents do matter: trajectories of change in externalizing and internalizing problems in early adolescence. *Child Dev,* 74**,** 578-94.

GALIANI, S., GERTLER, P. & SCHARGRODSKY, E. 2005. Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy,* 113**,** 83-120.

GALLAGHER, J. J. 1994. Teaching and Learning - New Models. *Annual Review of Psychology,* 45**,** 171-195.

GALLOP, K., ANDERSON, F. & AVRAHAM, B. 2015. Millennium Cohort Study, Fifth Sweep (MCS5) Teacher Survey;Technical Report. London: Centre for Longitudinal Studies, Institute for Education.

GALLOP, K., ROSE, N., WALLACE, E., WILLIAMS, R., CLEARY, A., THOMSON, A., BURSTON, K., FRERE-SMITH., DANGERFIELD, P. & TIETZ, S. 2013. Millennium Cohort Study Fifth Sweep (MCS 5); Technical Report. London: Ipsos MORI.

GAMORAN, A. & BERENDS, M. 1987. The Effects of Stratification in Secondary Schools: Synthesis of Survey and Ethnographic Research. *Review of Educational Research* 57**,** 415-435.

GARDINER, J. C., LUO, Z. H. & ROMAN, L. A. 2009. Fixed effects, random effects and GEE: What are the differences? *Statistics in Medicine,* 28**,** 221-239.

GARRATT, E. A., CHANDOLA, T., PURDAM, K. & WOOD, A. M. 2017. Income and Social Rank Influence UK Children's Behavioral Problems: A Longitudinal Analysis. *Child Development,* 88**,** 1302-1320.

GELMAN, A., HILL, J. & YAJIMA, M. 2012. Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness***,** 189–211.

GERSHOFF, T. E., MISTRY, R. S. & CROSBY, D. A. 2009. *Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy*, Oxford University Press.

GILLARD, D. 2009. Short and Fraught: the history of primary education in England. *FORUM,* 51**,** 143-164.

GLYMOUR, M. M., KAWACHI, I., JENCKS, C. S. & BERKMAN, L. F. 2008. Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology and Community Health,* 62**,** 532-537.

GOLDTHORPE, J. 2016. Social class mobility in modern Britain : changing structure, constant process *Journal of the British Academy***,** 89-111.

GOODMAN, A. & GREGG, P. 2010. Poorer children's educational attainment: how important are attitudes and behaviour? : Joseph Rowntree Foundation.

GOODMAN, A., LAMPING, D. L. & PLOUBIDIS, G. B. 2010. When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology,* 38**,** 1179-1191.

GOODMAN, A., PATEL, V. & LEON, D. A. 2008. Child mental health differences amongst ethnic groups in Britain: a systematic review. *Bmc Public Health,* 8.

GOODMAN, R. 1997. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry,* 38**,** 581-6.

GOODMAN, R. 1999. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *The Journal of Child Psychology and Psychiatry,* 40**,** 791-799.

GOODMAN, R. 2001. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry,* 40**,** 1337-45.

GORARD, S. & SIDDIQUI, N. 2018. Grammar schools in England: a new analysis of social segregation and academic outcomes. *British Journal of Sociology of Education,* 39**,** 909-924.

GOTTFREDSON, L. S. 1997. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography (Reprinted from The Wall Street Journal, 1994). *Intelligence,* 24**,** 13-23.

GREEN, A. 2003 Is UK Education Exceptionally Unequal? Evidence from the IALS and PISA Surveys. *FORUM,* 45**,** 67-70.

GREEN, A., PRESTON, J. & JANMAAT, J. G. 2006. *Education, equality and social cohesion: A comparative analysis. Basingstoke: ,* Basingstoke, Palgrave Macmillan.

GREEN, F. & KYNASTON, D. 2019. *Engines of Privilege-Britain's Private School Problem*, Bloomsbury Publishing.

GREENLAND, S. 1983. Tests for interaction in epidemiological studies: a review and a study of power. *Statistics and Medicine,* 2**,** 243-251.

GUMP, P. V. 1988. Citation Classic - Big School, Small School - High-School Size and Student Behavior. *Current Contents/Arts & Humanities***,** 14-14.

HALLAM, S., IRESON, J. & DAVIES, J. 2002. *Effective pupil grouping in the primary school -a practical guide.,* London, Fulton.

HALLAM, S., IRESON, J. & DAVIES, J. 2004a. Grouping practices in the primary school: what influences change? *British Educational Research Journal,* 30**,** 117-140.

HALLAM, S., IRESON, J. & DAVIES, J. 2004b. Primary pupils' experiences of different types of grouping in school. *British Educational Research Journal,* 30**,** 515-533.

HALLAM, S., IRESON, J., LISTER, V., CHAUDHURY, I. A. & DAVIES, J. 2003. Ability grouping practices in the primary school: a survey. *Educational Studies,* 29**,** 69-83.

HALLAM, S. & PARSONS, S. 2013a. The incidence and make up of ability grouped sets in the UK primary school. *Research Papers in Education,* 28**,** 393-420.

HALLAM, S. & PARSONS, S. 2013b. Prevalence of streaming in UK primary schools: evidence from the Millennium Cohort Study. *British Educational Research Journal,* 39**,** 514-544.

HALLINAN, M. & SORENSEN, A. 1985. Ability grouping and student friendships. *American Educational Research Journal***,** 485-499.

HANSEN, K., JOHNSON, J., CALDERWOOD, L., MOSTAFA, T., PLATT, L., ROSENBERG, R., SMITH, K. & TEAM, M. C. 2014. Millennium Cohort Study; A guide to the Datasets (Eighth Edition)-First, Second, Third, Fourth, Fifth Surveys. London: Institute of Education.

HANSEN, K. & KNEALE, D. 2013. Does How You Measure Income Make a Difference to Measuring Poverty? Evidence from the UK. *Social Indicators Research,* 110**,** 1119-1140.

HANUSHEK, E. A. & WOSSMANN, L. 2006. Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal,* 116**,** C63-C76.

HARDING, S., READ, U. M., MOLAODI, O. R., CASSIDY, A., MAYNARD, M. J., LENGUERRAND, E., ASTELL-BURT, T., TEYHAN, A., WHITROW, M. & ENAYAT, Z. E. 2015. The Determinants of young Adult Social well-being and Health (DASH) study: diversity, psychosocial determinants and health. *Social Psychiatry and Psychiatric Epidemiology,* 50**,** 1173-1188.

HARTAS, D. 2018. Setting for English and Maths: 11-year-olds' characteristics and teacher perceptions of school attitudes. *Research Papers in Education,* 33**,** 393-410.

HARTSHORNE, J. K., TENENBAUM, J. B. & PINKER, S. 2018. A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition,* 177**,** 263-277.

HEILMANN, A. 2013. *Neighbourhoods and children's social and cognitive development-pathways of effects* PhD, University College London.

HEILMANN, A., KELLY, Y., STAFFORD , M. & WATT, R. 2013. OP57 The Contribution of Neighbourhoods and Schools to Cognitive Test Performance at Age Seven - Findings from the UK Millennium Cohort Study. *Journal of Epidemiology and Community Health.*

HEMINGWAY, H. & MARMOT, M. 1999. Evidence based cardiology - Psychosocial factors in the aetiology and prognosis of coronary heart disease: systematic review of prospective cohort studies. *British Medical Journal,* 318**,** 1460-+.

HEO, M. & LEON, A. C. 2010. Sample Sizes Required to Detect Two-Way and Three-Way Interactions Involving Slope Differences in Mixed-Effects Linear Models. *Journal of Biopharmaceutical Statistics,* 20**,** 787-802.

HERNAN, M. A. & ROBINS, J. M. 2006. Estimating causal effects from epidemiological data. *J Epidemiol Community Health,* 60**,** 578-86.

HERTZMAN, C. 2012. Putting the concept of biological embedding in historical perspective. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 17160-17167.

HERTZMAN, C. & BOYCE, T. 2010. How Experience Gets Under the Skin to Create Gradients in Developmental Health. *Annual Review of Public Health, Vol 31,* 31**,** 329-347.

HILL, V. 2005. Through the Past Darkly: A Review of the British Ability Scales (Second Edition). *Child and Adolescent Mental Health,* 10 87-98.

HILL, W. D., MARIONI, R. E., MAGHZIAN, O., RITCHIE, S. J., HAGENAARS, S. P., MCINTOSH, A. M., GALE, C. R., DAVIES, G. & DEARY, I. J. 2019. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Molecular Psychiatry,* 24**,** 169-181.

HOPE, S., ROUGEAUX, E., DEIGHTON, J., LAW, C. & PEARCE, A. 2019. Associations between mental health competence and indicators of physical health and cognitive development in eleven year olds: findings from the UK Millennium Cohort Study. *Bmc Public Health,* 19.

HOUSE OF COMMONS. 2011. Behaviour and Discipline in Schools. .

HUANG, Y. & GATENBY, R. 2010. Millennium Cohort Study Sweep 4 Teacher Survey; Technical Report.

HUGHES, N., UNGAR, M., FAGAN, A., MURRAY, J., ATILOLA, O., NICHOLS, K., GARCIA, J. & KINNER, S. 2020. Health determinants of adolescent criminalisation. *Lancet Child Adolesc Health,* 4**,** 151-162.

HURRY, J., FLOURI, E. & SYLVA, K. 2018. Literacy Difficulties and Emotional and Behavior Disorders: Causes and Consequences. *Journal of Education for Students Placed at Risk,* 23**,** 259-279.

IRESON, J. & HALLAM, S. 2001. *Ability Grouping in Education; Issues and Context,* London, SAGE publications.

IRESON, J. & HALLAM, S. 2005. Pupils' liking for school: Ability grouping, self-concept and perceptions of teaching. *British Journal of Educational Psychology,* 75**,** 297-311.

IRESON, J. & HALLAM, S. 2009. Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction,* 19**,** 201-213.

IRESON, J., HALLAM, S. & HURLEY, C. 2005. What are the effects of ability grouping on GCSE attainment? *British Educational Research Journal,* 31**,** 443-458.
IRESON, J., HALLAM, S. & PLEWIS, I. 2001. Ability grouping in secondary schools: Effects on pupils' self-concepts. *British Journal of Educational Psychology,* 71**,** 315-326.
IRESON J., HALLAM S. & C, H. 2001. Ability grouping in Secondary School: Effects at Key Stage 4. London: The Nuffield Foundation.
IRWIN, G. L., SIDDIQI, A. & HERTZMAN, C. 2007. Early Child Development: A powerful Equalizer. Final Report.: WHO's Commission on Social Determinants of Health.
JAKOBSEN, I. S., HORWOOD, L. J. & FERGUSSON, D. M. 2012. Childhood Anxiety/Withdrawal, Adolescent Parent-Child Attachment and Later Risk of Depression and Anxiety Disorder. *Journal of Child and Family Studies,* 21**,** 303-310.
JEKIELEK, S. M. 1998. Parental Conflict, Marital Disruption and Children's Emotional Well-Being *Social Forces,* 76**,** 905-936.
JENSEN ARNETT, J. 2012. *Adolescence and Emerging Adulthood; A cultural-approach. Fifth Edition.*, Pearson.
JERRIM, J. & SIMS, S. 2018. The association between attending a grammar school and children'ssocio-emotional outcomes. New evidence from the Millennium Cohort Study.: UCL Institute of Education and Education Datalab
JERRIM, J. & VIGNOLES, A. 2013. Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 176**,** 887-906.
JERRIM, J., VIGNOLES, A., LINGAM, R. & FRIEND, A. 2015. The socio-economic gradient in children's reading skills and the role of genetics. *British Educational Research Journal* 41**,** 6-29.
JOHNSON, J., ROSENBERG, R., PLATT, L. & PARSONS , S. 2011. Millennium Cohort Study-Fourth Survey; A guide to the Teacher Survey Dataset (First Edition). London, Institute for Education: Centre for Longitudinal Studies.
JOSHI, H. 2014. Literature Review: Non-cognitive' skills: What are they and how can they be measured in the British cohort studies? *CLS working paper 2014/6.* London: UCL Institute of Education.
KAHN-LANG, A. & LANG, K. 2018. The promise and pitfalls of Differences-in-Differences: Reflections on '16 and pregnant' and other applications. Massachusets, Cambridge: National Bureau of Economic Research.
KAUTZ, T., HECKMAN, J. J., DIRIS, R., WEEL, B. T. & BORGHANS, L. 2014. FOSTERING AND MEASURING SKILLS: IMPROVING COGNITIVE AND NON-COGNITIVE SKILLS TO PROMOTE LIFETIME SUCCESS. *Working Paper 20749.* Cambridge, MA: NATIONAL BUREAU OF ECONOMIC RESEARCH
KELLY, Y., KELLY, J. & SACKER, A. 2013a. Changes in Bedtime Schedules and Behavioral Difficulties in 7 Year Old Children. *Pediatrics,* 132**,** E1184-E1193.
KELLY, Y., KELLY, J. & SACKER, A. 2013b. Time for bed: associations with cognitive performance in 7-year-old children: a longitudinal population-based study. *J Epidemiol Community Health,* 67**,** 926-31.
KELLY, Y., KELLY, J. & SACKER, A. 2013c. Time for bed: associations with cognitive performance in 7-year-old children: a longitudinal population-based study. *Journal of Epidemiology and Community Health,* 67**,** 926-931.
KELLY, Y., SACKER, A., DEL BONO, E., FRANCESCONI, M. & MARMOT, M. 2011. What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study. *Arch Dis Child,* 96**,** 832-7.
KERCKHOFF, A. 1986. Effects of ability grouping in British secondary schools. *American Sociological Review,* 51**,** 842-858

KESSLER, R. C., BERGLUND, P., DEMLER, O., JIN, R., MERIKANGAS, K. R. & WALTERS, E. E. 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry,* 62**,** 593-602.

KIDGER, J., ARAYA, R., DONOVAN, J. & GUNNELL, D. 2012. The effect of the school environment on the emotional health of adolescents: a systematic review. *Pediatrics,* 129**,** 925-49.

KIERNAN, K. E. & HUERTA, M. C. 2008. Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood. *Br J Sociol,* 59**,** 783-806.

KIERNAN, K. E. & MENSAH, F. K. 2009. Poverty, Maternal Depression, Family Status and Children's Cognitive and Behavioural Development in Early Childhood: A Longitudinal Study. *Journal of Social Policy,* 38**,** 569-588.

KIM, S. K. 2011. Consequences of Parental Divorce for Child Development. *Americal Sociological Review,* 76.

KITSON, N. & MERRY, R. 1997. *Teaching in the Primary School. A Learning Relationship (1st Edition),* London, Routledge.

KRAPOHL, E., 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 of the United States of America,* 111**,** 15273-15278.

KUH, D., BEN-SHLOMO, Y., LYNCH, J., HALLQVIST, J. & POWER, C. 2003. Life course epidemiology. *J Epidemiol Community Health,* 57**,** 778-83.

KUTNICK, P., HODGKINSON, S., SEBBA, J., HUMPHREYS, S., GALTON, M., STEWARD, S., BLATCHFORD, P. & BAINES, E. 2006. Pupil grouping strategies and practices at Key Stage 2 and 3: Case studies of 24 schools in England. (Research Report 796). London: Department for Education.

KUTNICK, P., SEBBA, J., BLATCHFORD, P., GALTON, M., THORPE, J., MACINTYRE, H. & BERDONDINI, L. 2005. The effects of pupil grouping: Literature review. *Research Report 688. .* London: Department for Education.

LACEY, C. 1970. *Hightown Grammar* Manchester, Manchester University Press.

LACEY, C. 1974. Destreaming in a "pressured" academic environment. *In:* EGGLESTON, J. (ed.) *Contemporary research in the sociology of education.* London.

LAREAU, A. 2011. Unequal Childhoods Class, Race, and Family Life SECOND EDITION WITH AN UPDATE A DECADE LATER Preface. *Unequal Childhoods: Class, Race, and Family Life, 2nd Edition***,** Ix-+.

LAVY, V., PASERMAN, M. D. & SCHLOSSER, A. 2012. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *Economic Journal,* 122**,** 208-237.

LAW, J., MCBEAN, K. & RUSH, R. 2011. Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage. *Int J Lang Commun Disord,* 46**,** 657-664.

LAW, J., RUSH, R., SCHOON, I. & PARSONS, S. 2009. Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes. *Journal of Speech Language and Hearing Research,* 52**,** 1401-1416.

LAZEAR, E. P. 2001. Educational production. *Quarterly Journal of Economics,* 116**,** 777-803.

LECKIE, G. 2009. The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 172**,** 537-554.

LEE, J. & CROLL, P. 1995. Streaming and Subject Specialism at Key-Stage-2 - a Survey in 2 Local-Authorities. *Educational Studies,* 21**,** 155-165.

LEVENTHAL, T., XUE, Y. & BROOKS-GUNN, J. 2006. Immigrant differences in school-age children's verbal trajectories: a look at four racial/ethnic groups. *Child Dev,* 77**,** 1359-74.
LEVIN, H. M. & BELFIELD, C. 2015. Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness,* 8 400-418.
LINCHEVSKI, L. & KUTSCHER, B. 1998. Tell me with whom you're learning and I'll tell you how much you've learned: Mixed ability versus same-ability grouping in mathematics. *Journal for Research in Mathematics Education* 533-544.
LINDSAY, G., DOCKRELL, J.E. AND MACKIE, L. 2008. Vulnerability to bullying in children with a history of specific speech and language difficulties. *European Journal of Special Needs Education,* 23**,** 1-16.
LIPSEY, M. W., PUZIO, K., YUNM C., HEBERT, M. A., STEINKA-FRY, K., COLE, M. W., ROBERTS, M., ANTHONY, K. S. & BUSICK, M. D. 2012. Translating the Statistical Representation of the Effects of Education Interventions Into More Readily Interpretable Forms. National Center for Special Education Research.
LIU, J. 2004. Childhood externalizing behavior: theory and implications. *J Child Adolesc Psychiatr Nurs,* 17**,** 93-103.
LOCKE, A., GINSBORG, J. & PEERS, I. 2002. Development and disadvantage: implications for the early years and beyond. *International Journal of Language & Communication Disorders,* 37**,** 3-15.
LOWE, R. 1997. The educational impact of the eugenics movement. *International Journal of Educational Research,* 27**,** 647–660.
LUMLEY, T., DIEHR, P., EMERSON, S. & CHEN, L. 2002. The importance of the normality assumption in large public health data sets. *Annual Review of Public Health,* 23**,** 151-169.
MACHIN, S., MCNALLY, S. & WYNESS, G. 2013. Educational attainment across the UK nations: performance, inequality and evidence. *Educational Research,* 55**,** 139-164.
MACINTYRE, H. & IRESON, J. 2002. Within-class Ability Grouping: Placement of pupils in groups and self-concept. *British Educational Research Journal,* 28**,** 249-263.
MAK, H. V. & FANCOURT, D. 2020. Longitudinal associations between reading for pleasure and child maladjustment: Results from a propensity score matching analysis. *Social Science and Medicine*.
MALDONADO, G. & GREENLAND, S. 2002. Estimating causal effects. *International Journal of Epidemiology,* 31**,** 422-429.
MARIANI, E., OZCAN, B. & GOISIS, A. 2017. Family Trajectories and Well-being of Children Born to Lone Mothers in the UK. *Eur J Popul,* 33**,** 185-215.
MARKHAM, W. A., YOUNG, R., SWEETING, H., WEST, P. & AVEYARD, P. 2012. Does school ethos explain the relationship between value-added education and teenage substance use? A cohort study. *Soc Sci Med,* 75**,** 69-76.
MARKS, R. 2016. *Ability-grouping in primary schools: Case studies and critical debates*, Critical Publishing
MARMOT, M. & WILKINSON, R. G. 2006. *Social Determinants of Health (Second Edition)*, Oxford University Press.
MARSH, W. M. 1984. Self-concept: The Application of a Frame of Reference Model to Explain Paradoxical Results. *Australian Journal of Education,* 28.
MAULANA, R., OPDENAKKER, M. C. & BOSKER, R. 2014. Teacher-student interpersonal relationships do change and affect academic motivation: A multilevel growth curve modelling. *British Journal of Educational Psychology,* 84**,** 459-482.
MAURIN, E. & MCNALLY, S. 2007. Widening access to grammar schools: The educational impact on Northern Ireland. . *Centrepiece,* 21**,** 26-27.
MAY, T. 2016. "Britain, the great meritocracy: Prime Minister's speech.".

MCCONNELL, B. & VERA-HERNÁNDEZ, M. 2015. Going beyond simple sample size calculations: a practitioner's guide. London: Institute of Fiscal Studies.
MCDOOL, E. 2018. Class Setting and Children's Non-Cognitive Outcomes. *Sheffield Economic Research Paper Series.* The University of Sheffield.
MCKINNON, B., HARPER, S., KAUFMAN, J. S. & BERGEVIN, Y. 2015. Removing user fees for facility-based delivery services: a difference-in-differences evaluation from ten sub-Saharan African countries. *Health Policy and Planning,* 30**,** 432-441.
MCLEOD, J. D. & SHANAHAN, J. M. 1993. Poverty, Parenting, and Children's Mental Health. *American Sociological Review,* 58**,** 351-366.
MCMANUS, J. E. 2010. *Social Identity and Ability Grouping in a Secondary School.* PhD, University of London.
MENSAH, F. K. & KIERNAN, K. E. 2010. Parents' mental health and children's cognitive and social development: families in England in the Millennium Cohort Study. *Soc Psychiatry Psychiatr Epidemiol,* 45**,** 1023-35.
MESCHI, E., MICKLEWRIGHT, J, VIGNOLES, J & GEOFF, L. 2010. The transitions between categories of special educational needs of pupils with Speech, Language and Communication Needs (SLCN) and Autism Spectrum Disorder (ASD) as they progress through the education system. Department of Education.
MESSER, J., MAUGHAN, B., QUINTON, D. & TAYLOR, A. 2004. Precursors and correlates of criminal behaviour in women. *Crim Behav Ment Health,* 14**,** 82-107.
MEYER, B. D. 1995. Natural and Quasi-Experiments in Economics. *Journal of Business & Economic Statistics,* 13**,** 151-161.
MIDOUHAS, E., KUANG, Y. & FLOURI, E. 2014. Neighbourhood human capital and the development of children's emotional and behavioural problems: The mediating role of parenting and schools. *Health & Place,* 27**,** 155-161.
MIHOUDAS, E. 2017. School poverty effects on trajectories of child behaviour: Do they depend on gender and ethnicity? *Health & Place,* 46**,** 281-292.
MILLER-JONES, D. 1989. Culture and testing. . *American Psychologist,* 44**,** 360-366.
MIROWSKY, J. & ROSS, C. E. 2002. Measurement for a human science. *Journal of Health and Social Behavior,* 43**,** 152-170.
MOFFITT, T. E., CASPI, A., HARRINGTON, H., MILNE, B. J., MELCHIOR, M., GOLDBERG, D. & POULTON, R. 2007. Generalized anxiety disorder and depression: childhood risk factors in a birth cohort followed to age 32. *Psychological Medicine,* 37**,** 441-452.
MOFFITT, T. E. & TEAM, E. R. S. 2002. Teen-aged mothers in contemporary Britain. *J Child Psychol Psychiatry,* 43**,** 727-42.
MORTIMER, J. T., ZHANG, F. L., HUSSEMANN, J. & WU, C. Y. 2014. Parental Economic Hardship and Children's Achievement Orientations. *Longit Life Course Stud,* 5**,** 105-128.
MOSS, G. & WASHBROOK, L. 2016. Understanding the Gender Gap in Literacy and Language Development. *Bristol Working Papers in Education* Bristol.
MOSTAFA, T. & ROSENBERG, R. 2013. Technical Report on Response in the Teacher Survey in MCS 4 (Age 7). London, Institute for Education: Centre for Longitudinal Studies.
MOSTAFA, T. & WIGGINS, R. 2015. The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies. *Longitudinal and Life Course Studies,* 6**,** 131-146.
MOULLIN, S., , WALDFOGEL, J. & WASHBROOK, E.
2014. Baby Bonds: Parenting, Attachment and a Secure Base for Children.
The Sutton Trust.
MUELLER, M. A. E., FLOURI, E. & KOKOSI, T. 2019. The role of the physical environment in adolescent mental health. *Health & Place,* 58.
MULLER, C. M. & HOFFMAN, V. 2016. Does Being Assigned to a Low School Track Negatively Affect Psychological Adjustment? Longitudinal Study in the First Year of Secondary School. *School Effectiveness and School Improvement 27,* 27**,** 95-115.

NEWMAN, B. M. & NEWMAN, P. R. 2015. *Development through life. A psychosocial approach.* , Cengage learning.
NICHOLSON, A., KUPER, H. & HEMINGWAY, H. 2006. Depression as an aetiologic and prognostic factor in coronary heart disease: A meta-analysis of 6608 events and 150217 participants in 63 observational studies. *Heart,* 92**,** A107-A107.
NOONAN, K., BURNS, R. & VIOLATO, M. 2018. Family income, maternal psychological distress and child socio-emotional behaviour: Longitudinal findings from the UK Millennium Cohort Study. *SSM Popul Health,* 4**,** 280-290.
NORBURY, C. F., GOOCH, D., WRAY, C., BAIRD, G., CHARMAN, T., SIMONOFF, E., VAMVAKAS, G. & PICKLES, A. 2016. The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *J Child Psychol Psychiatry,* 57**,** 1247-1257.
OAKES, J. M. & KAUFMAN, J. S. 2017. *Methods in Social Epidemiology*, Jossey-Bass.
OECD 2011. Strong Performers and Successful Reformers in Education; Lessons from PISA for the United States. OECD publishing.
OECD 2013. What makes schools successful? Resources, policies and practices.
OECD 2018a. Equity in Education: Breaking Down Barriers to Social Mobility, PISA. Paris.
OECD 2018b. Health at a Glance: Europe 2018; State of Health in the EU Cycle. Organisation for Economic Cooperation and Development.
OFFICE FOR NATIONAL STATISTICS. 2019. Schools, pupils and their characteristics: January 2019. Department for Education.
OKANO, L., JEON, L., CRANDALL, A. & RILEY, A. 2019. Differential effects of internalizing behaviors on academic functioning for girls versus boys: An analysis of developmental cascades from elementary to high school. *Dev Psychopathol***,** 1-14.
OLIVIER, E., ARCHAMBAULT, I. & DUPERE, V. 2018. Boys' and girls' latent profiles of behavior and social adjustment in school: Longitudinal links with later student behavioral engagement and academic achievement? *J Sch Psychol,* 69**,** 28-44.
PALAK, R., STYLES, B., WALKER, M., BRADSHAW, S., NELSON, J. & KETTLEWELL, K. 2018. Best Practice in Grouping Students Intervention B: Mixed Attainment Grouping; Pilot report and executive summary. Education Endowment Foundation.
PALAK , R., STYLES, B., WALKER, M., MORRISON, J., NELSON, J. & KETTLEWELL, K. 2018. Best Practice in Grouping Students Intervention A: Best Practice in Setting;Evaluation report and executive summary
Education Endowment Foundation.
PAPACHRISTOU, E., FLOURI, E., MIHOUDAS, E., LEWIS, G. & JOSHI, H. The Role of Primary School Composition in the Trajectories of Internalising and Externalising Problems across Childhood and Adolescence. *Journal of Abnormal Child Psychology*.
PARSONS, S., GREEN, F., PLOUBIDIS, G. B., SULLIVAN, A. & WIGGINS, R. D. 2017. The influence of private primary schooling on children's learning: Evidence from three generations of children living in the UK. *British Educational Research Journal,* 43**,** 823-847.
PARSONS, S. & HALLAM, S. 2014. The impact of streaming on attainment at age seven: evidence from the Millennium Cohort Study. *Oxford Review of Education,* 40**,** 567-589.
PATALAY, P. & FITZSIMONS, E. 2016. Correlates of Mental Illness and Wellbeing in Children: Are They the Same? Results From the UK Millennium Cohort Study. *Child and Adolescent Psychiatry,* 55**,** 771-783.
PATALAY, P. & FITZSIMONS, E. 2017. Mental ill-health among children of the new century: trends across childhood with a focus on age 14. London: Centre for Longitudinal Studies.
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**,** 542-557.

PLEWIS, I. 2007. The Millennium Cohort Study: Technical Report on Sampling 4th ed. London: Institute of Education.
PLOMIN, R. & SPINATH, F. M. 2002. Genetics and general cognitive ability (g). *Trends in Cognitive Sciences,* 6**,** 169-176.
PORTES, A. & MACLEOD, D. 1996. Educational progress of children of immigrants: The roles of class, ethnicity, and school context. *Sociology of Education,* 69**,** 255-275.
REAY, D. 2017. *Miseducation; Inequality, education and the working classes,* Great Britain, Policy Press.
REGGIO, I. & MORA, R. 2012. Treatment effect identification using alternative parallel assumptions. *UC3M Working papers. Economics we1233.* Universidad Carlos III de Madrid. Departamento de Economía.
RIGLIN, L., PETRIDES, K. V., FREDERICKSON, N. & RICE, F. 2014. The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence,* 37**,** 335-346.
RITTER, J. T. & HANCOCK, D. R. 2007. Exploring the relationship between certification sources, experience levels, and classroom management orientations of classroom teachers. *Teaching and Teacher Education,* 23**,** 1206-1216.
ROGERS, W. 1994. Regression standard errors in clustered samples
Stata Tech Bull. , 3**,** 19-23.
ROTHON, C., HEAD, J., CLARK, C., KLINEBERG, E., CATTELL, V. & STANSFELD, S. 2009. The impact of psychological distress on the educational achievement of adolescents at the end of compulsory education. *Social Psychiatry and Psychiatric Epidemiology,* 44**,** 421-427.
ROYSTON, P. & WHITE, I. R. 2011. Multiple Imputation by Chained Equations (MICE): Implementation in Stata. *Journal of Statistical Software,* 45**,** 1-20.
RUBIN, R. B. 1987. *Multiple Imputation for Nonresponse in Surveys,* New York, John Wiley and Sons, Inc.
RYAN, A., BURGESS, J. J. & DIMICK, J. 2015. Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. *Health Serv Res,* 50**,** 1211-35.
SACKER, A., CLARKE, P., WIGGINS, R. D. & BARTLEY, M. 2005. Social dynamics of health inequalities: a growth curve analysis of aging and self assessed health in the British household panel survey 1991-2001. *Journal of Epidemiology and Community Health,* 59**,** 495-501.
SADLER, K., VIZARD, T., FORD, T., GOODMAN, A., GOODMAN, R. & MCMANUS, S. 2018. Mental Health of Children and Young People in England, 2017: Trends and characteristics. NHS Digital.
SAVAGE, J. E., JANSEN, P. R., STRINGER, S., WATANABE, K., BRYOIS, J., DE LEEUW, C. A., NAGEL, M., AWASTHI, S., BARR, P. B., COLEMAN, J. R. I., GRASBY, K. L., HAMMERSCHLAG, A. R., KAMINSKI, J. A., KARLSSON, R., KRAPOHL, E., LAM, M., NYGAARD, M., REYNOLDS, C. A., TRAMPUSH, J. W., YOUNG, H., ZABANEH, D., HAGG, S., HANSELL, N. K., KARLSSON, I. K., LINNARSSON, S., MONTGOMERY, G. W., MUNOZ-MANCHADO, A. B., QUINLAN, E. B., SCHUMANN, G., SKENE, N. G., WEBB, B. T., WHITE, T., ARKING, D. E., AVRAMOPOULOS, D., BILDER, R. M., BITSIOS, P., BURDICK, K. E., CANNON, T. D., CHIBA-FALEK, O., CHRISTOFOROU, A., CIRULLI, E. T., CONGDON, E., CORVIN, A., DAVIES, G., DEARY, I. J., DEROSSE, P., DICKINSON, D., DJUROVIC, S., DONOHOE, G., CONLEY, E. D., ERIKSSON, J. G., ESPESETH, T., FREIMER, N. A., GIAKOUMAKI, S., GIEGLING, I., GILL, M., GLAHN, D. C., HARIRI, A. R., HATZIMANOLIS, A., KELLER, M. C., KNOWLES, E., KOLTAI, D., KONTE, B., LAHTI, J., LE HELLARD, S., LENCZ, T., LIEWALD, D. C., LONDON, E., LUNDERVOLD, A. J., MALHOTRA, A. K., MELLE, I., MORRIS, D., NEED, A. C., OLLIER, W., PALOTIE, A., PAYTON, A., PENDLETON, N., POLDRACK, R. A., RAIKKONEN, K., REINVANG, I., ROUSSOS, P., RUJESCU, D., SABB, F. W., SCULT, M. A., SMELAND, O. B., SMYRNIS, N., STARR, J. M., STEEN, V. M., STEFANIS, N. C., STRAUB, R. E., SUNDET, K., TIEMEIER, H., VOINESKOS, A. N., WEINBERGER, D. R., WIDEN, E., YU, J., ABECASIS, G., ANDREASSEN,

O. A., BREEN, G., CHRISTIANSEN, L., et al. 2018. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics,* 50**,** 912-+.
SAYAL, K., WASHBROOK, E. & PROPPER, C. 2015. Childhood Behavior Problems and Academic Outcomes in Adolescence: Longitudinal Population-Based Study. *Journal of the American Academy of Child and Adolescent Psychiatry,* 54**,** 360-368.
SCHOON, I., PARSONS, S., RUSH, R. & LAW, J. 2010. Children's language ability and psychosocial development: a 29-year follow-up study. *Pediatrics,* 126**,** e73-80.
SCHWARTZ, F. 1981. Supporting or subverting learning: Peer groups patterns in four tracked schools. *Anthropology and Education Quarterly,* 12**,** 99-121.
SCOURFIELD, J., MARTIN, N., LEWIS, G. & MCGUFFIN, P. 1999. Heritability of social cognitive skills in children and adolescents. *British Journal of Psychiatry,* 175**,** 559-564.
SEAMAN, S. R., WHITE, I. R., COPAS, A. J. & LI, L. 2012. Combining Multiple Imputation and Inverse-Probability Weighting. *Biometrics,* 68**,** 129-137.
SHARPLES, J., SLAVIN, R., CHAMBERS, B. & SHARP, C. 2011. Effective classroom strategies for closing the gap in educational achievement for children and young people living in poverty, including white working-class boys. London: Centre for Excellence and Outcomes in Children and Young People’s Services
SINDALL, K., STURGIS, P., STEELE, F., LECKIE, G. & FRENCH, R. 2019. A reassessment of socio-economic gradients in child cognitive development using growth mixture models. *Longitudinal and Life Course Studies,* 10**,** 283-305.
SKINNER, C. & WAKEFIELD, J. 2017. Introduction to the Design and Analysis of Complex Survey Data. *Statistical Science,* 32**,** 165-175.
SLAVIN, R. E. 1990. Achievement Effects of Ability Grouping in Secondary-Schools - a Best-Evidence Synthesis. *Review of Educational Research,* 60**,** 471-499.
SOCIAL MOBILITY COMMISSION 2017. Low-income pupils progress at secondary school.
SPIEGELHALTER, D. 2019. *The Art of Statistics: Learning from Data* Pelican Books.
STEELE, F., VIGNOLES, A. & JENKINS, A. 2007. The effect of school resources on pupil attainment: a multilevel simultaneous equation modelling approach. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 170**,** 801-824.
STEENBERGEN-HU, S., MAKEL, M. C. & OLSZEWSKI-KUBILIUS, P. 2016. What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K-12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses. *Review of Educational Research,* 86**,** 849-899.
STEINBERG, L. 2005a. Cognitive and affective development in adolescence. *Trends Cogn Sci,* 9**,** 69-74.
STEINBERG, L. 2005b. Cognitive and affective development in adolescence. *Trends in Cognitive Sciences,* 9**,** 69-74.
STRAND, S. 2014. School effects and ethnic, gender and socioeconomic gaps in educational achievement at age 11. *Oxford Review of Education,* 40**,** 223-245.
STROHSCHEIN, L. 2005. Parental Divorce and Child Mental Health Trajectories. *Journal of Marriage and Family,* 67**,** 1286-1300.
SUKHNANDAN, L. & LEE, B. 1998. Streaming, Setting and Grouping by Ability: A review of the literature. National Foundation for Educational Research.
SULLIVAN, A., MOULTON, V. & FITZSIMONS, E. 2017. The intergenerational transmission of vocabulary.
Centre for Longitudinal Studies Working paper 2017/14. London: UCL Institute of Education.
SULLIVAN, A., PARSONS, S., GREEN, F., WIGGINS, R., PLOUBIDIS, G. & HUYNH, T. 2018. Educational attainment in the short and long term: was there an advantage to attending faith, private and selective schools for pupils in the 1980s? *Oxford Review of Education*.

SUTCLIFFE, A. G. & DEROM, C. 2006. Follow-up of twins: Health, behaviour, speech, language outcomes and implications for parents. *Early Human Development,* 82**,** 379-386.

TERESHCHENKO, A., BECKY FRANCIS, B., ARCHER, L., HODGEN, J., MAZENOD, A., TAYLORM B, PEPPER, D. & TRAVERS, M.-C. 2019. Learners' attitudes to mixed-attainment grouping: examining the views of students of high, middle and low attainment. *Research Papers in Education,* 34**,** 425-444.

TOGA, A. W. & THOMPSON, P. M. 2005. Genetics of brain structure and intelligence. *Annual Review of Neuroscience,* 28**,** 1-23.

TOMBLIN, J. B., RECORDS, N. L., BUCKWALTER, P., ZHANG, X., SMITH, E. & O'BRIEN, M. 1997. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res,* 40**,** 1245-60.

TRAWICK-SMITH, J. 2008. *Early Childhood Development: A Multicultural Perspective ; 5 edition* Pearson.

VAN DEN BEDEM, N. P., DOCKRELL, J. E., VAN ALPHEN, P. M., KALICHARAN, S. V. & RIEFFE, C. 2018. Victimization, Bullying, and Emotional Competence: Longitudinal Associations in (Pre)Adolescents With and Without Developmental Language Disorder. *J Speech Lang Hear Res,* 61**,** 2028-2044.

VAN HOUTTE, M. 2006. School Type and Academic Culture: Evidence for the Differentiation–Polarization Theory *Journal of Curriculum Studies*

VAN HOUTTE, M. 2016. Lower-Track Students' Sense of Academic Futility: Selection or Effect? . *Journal of Sociology* 52**,** 874–889.

VAN HOUTTE, M. & STEVENS, P. A. J. 2008. Sense of futility. The missing link between track position and self-reported school misconduct. 40**,** 245-264.

VENKATAKRISHNAN, H. & WILIAM, D. 2003a. Tracking and mixed-ability grouping in secondary school mathematics classrooms: a case study. *British Educational Research Journal,* 29**,** 189-204.

VENKATAKRISHNAN, H. & WILIAM, D. 2003b. Tracking and mixed-ability grouping in secondary school mathematics classrooms: a case study *British Educational Research Journal* 29**,** 189-204.

VINCENT, C. & BALL, S. J. 2007. 'Making up' the middle-class child: Families, activities and class dispositions. *Sociology-the Journal of the British Sociological Association,* 41**,** 1061-1077.

VINER, R., WARD, J., CHEUNG, R., WOLFE , I. & HARGREAVES, D. 2018. Child health in 2030 in England: comparisons with other wealthy countries. *State of Child Health: short report series.* Royal College of Paediatrics and Child Health.

WALDFOGEL, J. & WASHBROOK, E. 2010. Low income and early cognitive development in the U.K. A Report for the Sutton Trust. London: The Sutton Trust.

WELLS, G. 2009. *The Meaning Makers: Children Learning Language and Using Language to Learn. Second Edition.,* Multilingual Matters.

WHARAM, J. F., LANDON, B. E., GALBRAITH, A. A., KLEINMAN, K. P., SOUMERAI, S. B. & ROSS-DEGNAN, D. 2007. Emergency department use and subsequent hospitalizations among members of a high-deductible health plan. *Jama-Journal of the American Medical Association,* 297**,** 1093-1102.

WHITBURN, J. 2001. Effective classroom organisation in primary schools: mathematics. *Oxford Review of Education,* 27**,** 411-428.

WHITE, I. R., ROYSTON, P. & WOOD, A. M. 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine,* 30**,** 377-399.

WHITE, J. 2006. *Intelligence, Destiny and Education: the ideological roots of intelligence testing.,* London, Routledge.

WHITESIDE, K. E., GOOCH, D. & NORBURY, C. F. 2017. English Language Proficiency and Early School Attainment Among Children Learning English as an Additional Language. *Child Development,* 88**,** 812-827.

WHITTY, G. & POWER, S. 2001. Devolution and choice in education - The research evidence to date. *Zeitschrift Fur Padagogik*, 99-118.
WILIAM, D. & BARTHOLOMEW, H. 2004. It’s not which school but which set you’re in that matters: The influence of ability grouping practices on student progress in mathematics. *British Educational Research Journal,* 2.
WILLIS, P. 1978. *Learning to Labour. How Working Class Kids Get Working Class Jobs.,* London, Routledge.
WING, C., SIMON, K. & BELLO-GOMEZ, R. A. 2018. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual Review of Public Health, Vol 39,* 39, 453-469.
WINSLER, A., CAVERLY, S. L., WILLSON-QUAYLE, A., CARLTON, M. P., HOWELL, C. & LONG, G. N. 2002. The social and behavioral ecology of mixed-age and same-age preschool classrooms: A natural experiment. *Journal of Applied Developmental Psychology,* 23, 305-330.
WOOLDRIDGE, J. M. 2013. *Introductory Econometrics: A Modern Approach,* Mason, OH., South-Western Cengage Learning.
ZAMMIT, S., LEWIS, G., DALMAN, C. & ALLEBECK, P. 2010. Examining interactions between risk factors for psychosis. *British Journal of Psychiatry,* 197, 207-211.
ZILANAWALA, A., KELLY, Y., SACKER, A. 2016. Ethnic Differences in Longitudinal Latent Verbal Profiles in the Millennium Cohort Study *Under review*.
ZILANAWALA, A., SACKER, A. & KELLY, Y. 2017. Longitudinal Latent Cognitive Profiles and Psychosocial Well-being in Early Adolescence. *J Adolesc Health,* 61, 493-500.
ZILANAWALA, A., SACKER, A., NAZROO, J. & KELLY, Y. 2015. Ethnic differences in children's socioemotional difficulties: Findings from the Millennium Cohort Study. *Social Science & Medicine,* 134, 95-106.

Appendix 1

Table A1. 1: Difference in sample composition between the entire eligible MCS sample at MCS 4 and MCS 5 and the eligible MCS Cohort members for whom there were teacher responses at MCS 4 and MCS 5

Variables	Categories	Eligible MCS sample at MCS4 & MCS5	MCS sample at both MCS4 & MCS5	% Difference
MCS strata	England-Advantaged	54.1%	59.8%	5.7%
	England - Disadvantaged	34.3%	31.3%	-3.0%
	England - Ethnic	6.5%	4.2%	-2.3%
	Wales - Advantaged	2.6%	2.5%	-0.1%
	Wales - Disadvantaged	2.5%	2.2%	-0.3%
Labour market status	Two-parent family, both work	48.0%	52.0%	4%
	Two-parent family, one works	20.2%	19.8%	-0.4%
	Both not in work	6.0%	4.8%	-1.2%
	One-parent family, in work	14.3%	14.5%	0.2%
	One-parent family, not in work	11.5%	9.0%	-2.5%
Family income quintiles	Highest incomes (Top quantile)	21.8%	25.4%	3.5%
	2nd	21.0%	24.3%	3.3%

Variables	Categories	Eligible MCS sample at MCS4 & MCS5	MCS sample at both MCS4 & MCS5	% Difference
	3rd	19.7%	19.3%	-0.4%
	4th	18.8%	17.6%	-1.3%
	Lowest incomes	18.6%	13.4%	-5.1%
Ethnic group	White	84.4%	87.8%	3.4%
	Mixed	3.8%	3.4%	-0.4%
	Indian	2.2%	2.0%	-0.2%
	Pakistani	3.5%	2.5%	-0.9%
	Bangladeshi	1.2%	0.6%	-0.6%
	Black Caribbean	1.1%	0.8%	-0.3%
	Black African	2.1%	1.5%	-0.6%
	Other ethnic group	1.8%	1.4%	-0.5%
Language spoken at home	English only	90.5%	93.1%	2.6%
	Mostly English-sometimes other	3.7%	2.9%	-0.8%
	About half English, half other	2.6%	2.0%	-0.6%
	Mostly other	3.2%	2.0%	-1.2%
Maternal psychological distress	No			0.0%
	Yes	6.1%	4.9%	-1.2%
Very interested		58.1%	61.9%	3.8%
Main respondent's age at interview		39.0	40.2	
Number of observations		5,106	4,469	

Table A1. 2: Logistic regression of response to both MCS 4 and MCS 5 teacher surveys in England and Wales.

		With survey design +attrition weights	
		OR	**SE**
Variables	**Categories**		
MCS strata	England Advantaged (Ref.)		
	England - Disadvantaged	0.87	0.08
	England - Ethnic	*0.67**	*0.11*
	Wales - Advantaged	0.75	0.12
	Wales - Disadvantaged	0.85	0.09
	Scotland - Advantaged		
Labour market status	Two-parent family, both work (Ref.)		
	Two-parent family, one works	*1.22**	*0.10*
	Both not in work	1.20	0.22
	One-parent family, in work	1.31**	0.13
	One-parent family, not in work	1.19	0.17
Family income quintiles	Highest incomes (Top quantile)		
	2nd	1.07	0.08
	3rd	0.72*	0.07
	4th	0.70*	0.09
	Lowest incomes	*0.51****	*0.08*
Ethnic group	White (Ref.)		

		With survey design +attrition weights	
		OR	**SE**
	Mixed	0.87	0.13
	Indian	1.11	0.24
	Pakistani	0.99	0.24
	Bangladeshi	*0.56**	*0.14*
	Black Caribbean	0.75	0.28
	Black African	0.69	0.14
	Other ethnic group	0.74	0.18
Language spoken at home	English only (Ref.)		
	Mostly English-sometimes other	0.90	0.13
	About half English, half other	0.73	0.17
	Mostly other	0.70	0.15
	Only other		
Maternal psychological distress	No (Ref.)		
	Yes	0.84	0.11
Maternal age at birth	(older vs. younger)	1.02**	0.01
Observations (N)		9,575	
Pseudo R2		0.03	

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A1. 3 Variables used for the analyses of section 4.3.1.3

Variable name	Description	MCS Variable(s) used
Outcomes		
Stream group at MCS 4	0=Not in stream, 1=Top stream, 2=Middle stream, 3=Bottom stream	stream
Stream group at MCS 5	0=Not in stream, 1=Top stream, 2=Middle stream, 3=Bottom stream	STREAM
Literacy set at MCS 4	0=Not in set for literacy, 1=Top set for literacy, 2=Middle set for literacy, 3=Bottom set for literacy	litset
Literacy set at MCS 5	0=Not in set for literacy, 1=Top set for literacy, 2=Middle set for literacy, 3=Bottom set for literacy	LITSET
Maths set at MCS 4	0=Not in set for maths, 1=Top set for maths, 2=Middle set for maths, 3=Bottom set for maths	numset

Variable name	Description	MCS Variable(s) used
Maths set at MCS 5	0=Not in set for maths, 1=Top set for maths, 2=Middle set for maths, 3=Bottom set for maths	NUMSET
Explanatory variables		
Child's gender	0=Male, 1=Female	ECQ37X00
Ethnic group	0=White UK, 1=Mixed, 2=Indian, 3=Pakistani, 4=Bangladeshi, 5=Black Caribbean, 6=Black African, 7=Other	EDC08E00
Time of the year born	0=Autumn 2000, 1=Winter 2000/1, 2=Spring 2001, 3=Summer/Autumn/Winter 2001	TIMEYEARG
Equivalised family income at MCS 5	0=Highest quintile, 1=2nd quintile, 2=3rd quintile, 3=4th quintile, 4=Lowest quintile	EOECDUK0
Highest parental qualifications	0=NVQ 4/5, 1=NVQ3=1, NVQ2=2 NVQ1=3, 3=Overseas qualifications, 4=No qualifications	EDDNVQ001 & EDDNVQ002

Variable name	Description	MCS Variable(s) used
Family structure at MCS 4	0=Two-parent family, 1=Single-parent family	DDHTYS00
Family structure at MCS 5	0=Two-parent family, 1=Single-parent family	EDHTYS00
BAS naming vocabulary (age 5)	Scores range from 20 to 80. A higher score indicates a higher verbal cognitive ability	cdnvtscr
Parent-reported externalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	DDHYPEA0+DDCONDA0

Variable name	Description	MCS Variable(s) used
Parent-reported externalising SDQ at MCS 5	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	EDHYPEA00+EDCOND00
Child likes school (MCS 4)	0=A lot, 1=A bit, 2=Does not like it	dcsc0021
Child likes school (MCS 5)	0=A lot, 1=A bit, 2=Does not like it	ECQ29X00
Child unhappy at school (MCS 4)	0=All the time, 1=Most/some of the time, 2=Never	dcsc0032
Child unhappy school (MCS 5)	0=All the time, 1=Most/some of the time, 2=Never	ECQ37X00
Child tries her/his best at school (MCS 4)	0=All the time, 1=Most/some of the time, 2=Never	dcsc0032
Child tries her/his best at school (MCS 5)	0=All the time, 1=Most/some of the time, 2=Never	ECQ35X00

Appendix 2

Table A2. 1: Weighted logistic regression of teacher item non-response on variables related to transition to attainment grouping between 7 and 11 years.

	OR	S.E
MCS strata (Ref. England advantaged)		
England - Disadvantaged	1.02	0.07
England - Ethnic	1.69****	0.16
Wales - Advantaged	0.55***	0.08
Wales - Disadvantaged	0.68***	0.07
School type (Ref. State-maintained)		
Fee-paying school	1.00	0.14
Faith school	0.84**	0.06

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 2: Weighted regression of missing information on child BAS cognitive assessments at MCS 3, MCS 4 and MCS 5.

	Verbal skills at 5		**Verbal skills at 7**		**Verbal skills at 11**	
	OR	**SE**	**OR**	**SE**	**OR**	**SE**
MCS strata (Ref. England-Advantaged)						
England - Disadvantaged	1.61	0.55	1.76***	0.24	0.89	0.28
England - Ethnic	1.94	0.78	1.96***	0.33	0.62	0.29
Wales - Advantaged	0.96	0.72	3.57***	0.70	0.88	0.54
Wales - Disadvantaged	2.63*	1.01	3.27***	0.51	0.98	0.41
Highest incomes (Ref. Top quantile)						
2nd quintile	0.96	0.48	1.17	0.21	0.96	0.42
3rd quintile	0.91	0.46	1.07	0.20	1.10	0.48
4th quintile	2.62*	1.11	2.01***	0.34	1.81	0.73
Lowest incomes	2.15	0.94	2.47***	0.41	1.56	0.67

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A2. 3: Variable labels and definitions for empirical analysis in chapter 5.

Variable name	Description	MCS Variable(s) used
Key Explanatory variable		
Attainment grouping at MCS 4	The 3 variables were used to derive the attainment grouping status at age 7 where 0=Not grouped by attainment, 1=Grouped by attainment (either for streaming or literacy setting or maths setting)	streamg, litsetg, numsetg
Attainment grouping at MCS 5	The 3 variables were used to derive the attainment grouping status at age 11 where 0=Not grouped by attainment, 1=Grouped by attainment (either for streaming or literacy setting or maths setting)	STREAMG, LITSETG, NUMSETG
Outcome(s)		
BAS word reading (MCS 4)	The word reading score was recoded in order to have a mean of 56.5 and a standard deviation of 10. Scores range from 25 to 75. A higher score indicates higher reading ability	DCWRSD00

BAS verbal similarities (MCS 5)	Scores range from 20 to 80. A higher score indicates higher verbal cognitive ability.	EVSTSCO
Effect modifiers		
BAS naming vocabulary (MCS 3)	Scores range from 20 to 80. A higher score indicates a higher verbal cognitive ability	cdnvtscr
Family income (MCS 5)	Equivalised family income quantiles where 0=Highest incomes, 1=2nd quantile, 2=3rd quantile, 3=4th quantile, 4=Lowest incomes	EOECDUK0
Sensitivity analyses		
School moves since MCS 4	0=MCS participant did not change school between MCS 4 and MCS 5	EPSAMS00
	1=MCS participant changed school between MCS 4and MCS 5	
Relationship between streaming and literacy setting at MCS 4	0=MCS participant was neither streamed nor set for literacy, 1=CM was either streamed or set for literacy	strlitg
Relationship between streaming and literacy setting at MCS 5	0=MCS participant was neither streamed nor set for literacy, 1=MCS participant was either streamed or set for literacy	STREAMLITG
Stream group at MCS 4	0=Not in stream	stream

	1=Top stream 2=Middle stream 3=Bottom stream	
Literacy set at MCS 4	0=Not in set for literacy 1=Top set for literacy 2=Middle set for literacy 3=Bottom set for literacy	litset
Numeracy set at MCS 4	0=Not in set for maths 1=Top set for maths 2=Middle set for maths 3=Bottom set for maths	numset
Stream group at MCS 5	0=Not in stream 1=Top stream 2=Middle stream 3=Bottom stream	STREAM
Literacy set at MCS 5	0=Not in set for literacy 1=Top set for literacy 2=Middle set for literacy	LITSET

	3=Bottom set for literacy	
Numeracy set at MCS 5	0=Not in set for maths 1=Top set for maths 2=Middle set for maths 3=Bottom set for maths	NUMSET
Number of classes in child's school year at MCS 4	Range 1 to 18	dq2524
Number of classes in child's school year at MCS 5	Range 1 to 10	EQ35
Number of children within child's classroom at MCS 4	Range 1 to 48	dq2511
Number of children within child's classroom at MCS 5	Range 1 to 64	EQ33
Number of children with English as additional language in child's classroom at MCS 4	Range 0 to 39	dq2524

Number of children with English as additional language in child's classroom at MCS 5	Range 0 to 88	EQ38
Child's class contain a mixed year group at MCS 4	0=No	dq2513
	1=Yes	
Child's class contain a mixed year group at MCS 5	0=No	EQ34
	1=Yes	

Table A2. 4: Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores (adjusted for school related characteristics)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.35***	0.26	0.84	1.85
Attainment grouping * MCS5	-0.03	0.41	-0.84	0.77
School size (number of classes in child's school year)	-0.08	0.05	-0.17	0.02
Number of children within child's classroom	0.11	0.20	-0.28	0.51
Number of EAL children in child's classroom	-0.03	0.04	-0.10	0.04
Child in mixed year classroom	-0.27	0.44	-1.13	0.60

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A2. 5: *Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by family income (adjusted for school related characteristics)*

VARIABLES	B	(SE)	95% CI	95% CI
Change from MCS 4 to MCS 5	0.90	0.47	-0.03	1.83
Attainment grouping * MCS5	-0.06	0.67	-1.37	1.25
Highest incomes * MCS5 (Ref.)				
2nd quantile * MCS5	0.02	0.83	-1.61	1.65
3rd quantile * MCS5	1.09	0.79	-0.45	2.64
4th quantile * MCS5	2.11*	0.87	0.41	3.82
Lowest incomes * MCS5	0.91	0.83	-0.71	2.53
Attainment grouping * Highest incomes * MCS5 (Ref.)				
Attainment grouping * 2nd quantile * MCS5	0.83	1.19	-1.51	3.17
Attainment grouping * 3rdquantile MCS5	0.33	1.09	-1.80	2.46
Attainment grouping* 4th quantile * MCS5	-1.20	1.23	-3.60	1.21

VARIABLES	B	(SE)	95% CI	95% CI
Attainment grouping * Lowest incomes * MCS5	-1.21	1.14	-3.44	1.03
School size (number of classes in child's school year)	0.26	0.21	-0.16	0.67
Number of children within child’s classroom	0.00	0.05	-0.10	0.10
Number of EAL children in child's classroom	-0.05	0.04	-0.12	0.02
Child in mixed year classroom	-0.31	0.35	-1.01	0.38

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 6: Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by initial verbal skills (adjusted for school related characteristics)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	2.44*	1.23	0.04	4.85
Attainment grouping * MCS5	-2.82	1.74	-6.25	0.61
Baseline verbal cognitive skills* MCS5	-0.01	0.02	-0.06	0.03
Baseline verbal cognitive skills* Attainment grouping * MCS5	0.05	0.03	-0.01	0.11
School size (number of classes in child's school year)	0.22	0.21	-0.20	0.64
Number of children within child's classroom	-0.02	0.05	-0.12	0.07
Number of EAL children in child's classroom	-0.04	0.04	-0.11	0.03
Child in mixed year classroom	-0.29	0.35	-0.98	0.41

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A2. 7:Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by initial verbal skills (adjusted for school moves)

VARIABLES	B	(SE)	95% CI	95% CI
Change from MCS 4 to MCS 5	1.66*	0.25	1.17	2.15
Attainment grouping * MCS 5	-0.19	0.35	-0.88	0.50
School change between MCS4 and MCS 5	-0.03	0.36	-0.74	0.67

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 8:Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by family income (adjusted for school moves)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.71	1.06	-0.37	3.78
Attainment grouping * MCS5	-0.89	1.26	-3.35	1.57
Highest incomes * MCS5 (Ref.)				
2nd quantile * MCS5	-1.58	1.66	-4.84	1.69
3rd quantile * MCS5	1.15	1.54	-1.86	4.16
4th quantile * MCS5	2.63	1.56	-0.43	5.68
Lowest incomes * MCS5	-1.18	1.61	-4.34	1.98
Attainment grouping * Highest incomes * MCS5 (Ref.)				
Attainment grouping * 2nd quantile * MCS5	1.31	2.01	-2.63	5.25
Attainment grouping * 3rdquantile MCS5	-0.17	1.88	-3.86	3.52
Attainment grouping* 4th quantile * MCS5	-1.88	2.01	-5.82	2.06
Attainment grouping * Lowest incomes * MCS5	1.95	2.17	-2.31	6.20
School change between MCS4 and MCS 5	0.03	0.63	-1.20	1.26

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A2. 9:Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by family income (adjusted for school moves)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	2.45	1.23	0.03	4.87
Attainment grouping * MCS5	-2.92	1.74	-6.36	0.51
Baseline verbal cognitive skills* MCS5	-0.01	0.02	-0.06	0.03
Baseline verbal cognitive skills* Attainment grouping * MCS5	0.05	0.03	-0.01	0.11
School change between MCS4 and MCS 5	-0.05	0.36	-0.76	0.65

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 10:Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores (considering transition to streaming/literacy setting only)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.77***	0.3	1.18	2.35
Attainment grouping * MCS5	-0.66	0.45	-1.56	0.23

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** $p<0.001$, ** <0.01, * <0.05

Table A2. 11:*Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by family income (considering transition to streaming/literacy setting only)*

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.54*	0.73	0.11	2.97
Attainment grouping * MCS5	-0.58	1.00	-2.53	1.37
Highest incomes * MCS5 (Ref.)				
2nd quantile * MCS5	-1.49	1.13	-3.71	0.74
3rd quantile * MCS5	0.80	1.08	-1.32	2.92
4th quantile * MCS5	2.43*	1.20	0.07	4.77
Lowest incomes * MCS5	0.01	1.26	-2.46	2.47
Attainment grouping * Highest incomes * MCS5 (Ref.)				
Attainment grouping * 2nd quantile * MCS5	1.94	1.64	-1.26	5.15
Attainment grouping * 3rdquantile MCS5	0.41	1.60	-2.73	3.55
Attainment grouping* 4th quantile * MCS5	-2.84	1.87	-6.49	0.82
Attainment grouping * Lowest incomes * MCS5	-1.03	2.02	-4.99	2.93

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** p<0.001, ** <0.01, * <0.05*

Table A2. 12:*Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by verbal skills at age 5 (considering transition to streaming/literacy setting only)*

VARIABLES	B	(SE)	95% CI	95% CI
Change from MCS 4 to MCS 5	3.67*	1.72	0.29	7.04
Attainment grouping * MCS5	-6.53*	2.54	-11.5	-1.55
Baseline verbal cognitive skills* MCS5	-0.03	0.03	-0.09	0.02
Baseline verbal cognitive skills* Attainment grouping * MCS5	0.10*	0.04	0.02	0.19

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 13 Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by family income (adjusting for stream/set group allocation)

VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	1.37	0.99	-0.57	3.32
Attainment grouping * MCS5	-0.16	1.81	-3.71	3.39
Highest family incomes * MCS5				
2nd quantile * MCS5	0.95	1.40	-1.79	3.70
3rd quantile * MCS5	1.87	1.49	-1.05	4.78
4th quantile * MCS5	2.78	1.78	-0.71	6.28
Lowest family incomes * MCS5	3.61	2.07	-0.45	7.66
Highest family incomes * Attainment grouping * MCS5 (Ref.)				
2nd quintile * Attainment grouping * MCS5	0.24	1.89	-3.48	3.95
3rd quintile * Attainment grouping * MCS5	-1.21	1.80	-4.74	2.33
4th quintile * Attainment grouping * MCS5	-3.19	1.97	-7.05	0.68
Lowest family incomes * Attainment grouping * MCS5 (Ref.)	-0.16	2.13	-4.33	4.02
Transition to stream groups (Ref. Mixed)				
Top	0.04	0.92	-1.77	1.84
Middle	1.09	1.08	-1.03	3.21
Bottom	1.20	1.91	-2.54	4.94
Transition to literacy set groups (Ref. Mixed)				
Top	-1.99**	0.76	-3.49	-0.50
Middle	1.25	0.91	-0.54	3.05

VARIABLES	B	(SE)	95% CI	95% CI
Bottom	3.81**	1.36	1.13	6.48
Transition to maths set groups (Ref. Mixed)				
Top	-0.23	1.38	-2.93	2.48
Middle	0.38	1.51	-2.59	3.34
Bottom	-0.05	1.66	-3.30	3.20

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A2. 14:Sensitivity analysis: The association between transition to attainment grouping between age 7 and 11 years and changes in children's verbal cognitive scores by verbal skills at age 5 (adjusting for stream/set group allocation)

	Multiple imputation (n=4,166)			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	3.39	2.19	-0.90	7.68
Attainment grouping * MCS5	-7.19*	3.11	-13.30	-1.08
Baseline verbal cognitive skills* MCS5	-0.03	0.04	-0.10	0.05
Baseline verbal cognitive skills* Attainment grouping * MCS5	0.12*	0.05	0.02	0.21
Transition to stream groups				
Top	-0.03	0.90	-1.79	1.73
Middle	1.24	1.11	-0.93	3.41
Bottom	3.20	1.89	-0.50	6.90
Transition to literacy set groups				
Top	-2.12*	0.76	-3.61	-0.62
Middle	1.37	0.94	-0.47	3.22
Bottom	3.30*	1.34	0.68	5.93
Transition to maths set groups				
Top	-0.66	1.340	-3.40	2.08
Middle	-0.08	1.52	-3.06	2.90

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Appendix 3

Table A3. 1: Weighted logistic regression of non-response to parent reported SDQ externalising subscale at MCS 4, MCS 5 and MCS 6, by MCS strata and family income.

	Parent-reported externalising scores at MCS 4 (age 7)		**Parent-reported externalising scores at MCS 5 (age 11)**		**Parent-reported externalising scores at MCS 6 (age 14)**	
	OR	**SE**	**OR**	**SE**	**OR**	**SE**
MCS strata (Ref. England advantaged)	1.78*	0.49	1.25	0.31	0.65	0.14
England - Disadvantaged	5.67***	1.54	4.65***	1.13	1.09	0.28
England - Ethnic	0.58	0.43	1.40	0.59	0.29*	0.17
Wales - Advantaged	0.48	0.24	1.55	0.47	1.44	0.34
Wales - Disadvantaged	0.46	0.34	0.96	0.40	0.60	0.21
Scotland - Advantaged	0.54	0.34	1.45	0.54	0.43	0.20
Scotland - Disadvantaged	0.52	0.39	1.22	0.55	0.67	0.27
Northern Ireland - Advantaged	2.54**	0.86	1.31	0.46	1.47	0.40
Northern Ireland - Disadvantaged						
Highest incomes (Ref. Top quantile)						
2nd quantile	4.92*	3.11	0.69	0.20	0.74	0.17
3rd quantile	4.52*	2.85	0.77	0.22	1.15	0.25
4th quantile	13.63***	8.18	1.76*	0.46	1.18	0.29
Lowest incomes	15.15***	9.09	2.64***	0.68	1.53	0.38

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A3. 2: Weighted logistic regression of non-response to parent reported SDQ internalising subscale at MCS 4, MCS 5 and MCS 6 by MCS stratum and family income.

	Parent-reported internalising scores at MCS 4 (age 7)		**Parent-reported internalising scores at MCS 5(age 11)**		**Parent-reported internalising scores at MCS 6 (age 14)**	
	OR	**SE**	**OR**	**SE**	**OR**	**SE**
MCS strata (Ref. England advantaged)						
England - Disadvantaged	1.69*	0.45	1.43	0.31	0.63*	0.14
England - Ethnic	5.30***	1.40	4.16***	0.93	0.97	0.26
Wales - Advantaged	0.81	0.50	1.33	0.49	0.28*	0.17
Wales - Disadvantaged	0.64	0.28	1.44	0.40	1.44	0.34
Scotland - Advantaged	0.21	0.21	1.24	0.41	0.59	0.21
Scotland - Disadvantaged	0.84	0.42	1.42	0.47	0.43	0.20
Northern Ireland - Advantaged	-	-	1.20	0.47	0.67	0.27
Northern Ireland - Disadvantaged	2.24*	0.76	1.09	0.37	1.47	0.40
Highest incomes (Ref. Top quantile)						
2nd quantile	2.76	1.43	1.01	0.23	0.72	0.17
3rd quantile	3.20*	1.62	0.79	0.20	1.11	0.25
4th quantile	8.24***	3.92	1.41	0.33	1.14	0.28
Lowest incomes	8.93***	4.25	2.06**	0.48	1.48	0.37

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A3. 3: Logistic regression of non-response to teacher reported SDQ externalising subscale at MCS 4 and MCS 5, by MCS strata and primary school type (England and Wales).

	Teacher-reported externalising scores at MCS 4 (age 7)		**Teacher-reported externalising scores at MCS 5 (age 11)**	
	OR	**SE**	**OR**	**SE**
MCS strata (Ref. England advantaged)				
England - Disadvantaged	1.28***	0.08	0.72	0.14
England - Ethnic	2.40***	0.20	0.57	0.17
Wales - Advantaged	0.97	0.12	0.39	0.20
Wales - Disadvantaged	1.30**	0.12	0.82	0.23
School type (Ref. State-Maintained)				
Fee-paying	0.63**	0.092	1.07	0.40
Faith school	0.90	0.056	0.71	0.15

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A3. 4: Logistic regression of non-response to teacher reported SDQ internalising subscale at MCS sweeps 4 and 5, by MCS strata and primary school type (England and Wales).

	Teacher-reported internalising scores at MCS 4 (age 7)		**Teacher-reported internalising scores at MCS 5 (age 11)**	
	OR	**SE**	**OR**	**SE**
MCS strata (Ref. England advantaged)				
England - Disadvantaged	1.28***	0.08	0.94	0.16
England - Ethnic	2.41***	0.20	0.69	0.18
Wales - Advantaged	0.97	0.12	0.42	0.20
Wales - Disadvantaged	1.30**	0.12	0.65	0.18
School type (Ref. State-maintained)				
Fee-paying	0.63**	0.09	0.94	0.33
Faith school	0.89	0.06	0.58**	0.11

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A3. 5: Labels and definitions for empirical analysis in chapter 6.

Variable name	**Description**	**MCS Variable(s) used**
Key Explanatory variable		
Attainment grouping at MCS 47	The 3 variables were used to derive the attainment grouping status at age 7 where 0=Not grouped by attainment, 1=Grouped by attainment (either for streaming or literacy setting or maths setting)	streamg, litsetg, numsetg
Attainment grouping at MCS 5	The 3 variables were used to derive the attainment grouping status at age 11 where 0=Not grouped by attainment, 1=Grouped by attainment (either for streaming or literacy setting or maths setting)	STREAMG, LITSETG, NUMSETG
Outcome(s)		
Parent-reported externalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	DDHYPEA0+DDCONDA0

Variable name	Description	MCS Variable(s) used
Parent-reported externalising SDQ at MCS 5	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	EDHYPEA00+EDCOND00
Parent-reported internalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	DDEMOTA0+DDPEERA0
Parent-reported internalising SDQ at MCS 5	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	EDEMOT00+EDPEER00

Variable name	Description	MCS Variable(s) used
Teacher-reported externalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	ddhypea0+ddconda0
Teacher-reported externalising SDQ at MCS 5	An externalising SDQ score was derives from these 10 variables. Variables with an asterisk were reverse coded. A higher score indicates higher externalising problem behaviour symptoms.	EQ5E+EQ5L+EQ5R+EQ5V+EQ5G*+EQ5B+EQ5J+EQ5O+EQ5U*EQ5Y*
Teacher-reported internalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	dpeer00+demot00
Teacher-reported internalising SDQ at MCS 5	An internalising SDQ score was derives from these 10 variables. Variables with an asterisk were	EQ5C, EQ5H, EQ5M, EQ5P, EQ5X, EQ5K*, EQ5N*, EQ5F, EQ5S, EQ5W

Variable name	Description	MCS Variable(s) used
	reverse coded. A higher score indicates higher internalising problem behaviour symptoms.	
Effect modifiers		
Family income	Equivalised family income quintiles where 0=Highest incomes, 1=2nd quantile, 2=3rd quantile, 3=4th quantile, 4=Lowest incomes	EOECDUK0
BAS naming vocabulary (MCS 3)	Scores range from 20 to 80. A higher score indicates a higher verbal cognitive ability	cdnvtscr
Sensitivity analyses		
School moves since MCS 4 (MCS 5)	0=MCS participant did not change school between MCS 4 and MCS 5	EPSAMS00
	1=MCS participant changed school between MCS 4and MCS 5	
Relationship between streaming and literacy setting at MCS 4	0=MCS participant was neither streamed nor set for literacy, 1=CM was either streamed or set for literacy	strlitg

Variable name	Description	MCS Variable(s) used
Relationship between streaming and literacy setting at MCS 5	0=MCS participant was neither streamed nor set for literacy, 1=MCS participant was either streamed or set for literacy	STREAMLITG
Stream group at MCS 4	0=Not in stream	stream
	1=Top stream	
	2=Middle stream	
	3=Bottom stream	
Literacy set at MCS 4	0=Not in set for literacy	litset
	1=Top set for literacy	
	2=Middle set for literacy	
	3=Bottom set for literacy	
Numeracy set at MCS 4	0=Not in set for maths	numset
	1=Top set for maths	
	2=Middle set for maths	
	3=Bottom set for maths	
Stream group at MCS 5	0=Not in stream	STREAM

Variable name	**Description**	**MCS Variable(s) used**
	1=Top stream	
	2=Middle stream	
	3=Bottom stream	
Literacy set at MCS 5	0=Not in set for literacy	LITSET
	1=Top set for literacy	
	2=Middle set for literacy	
	3=Bottom set for literacy	
Numeracy set at MCS 5	0=Not in set for maths	NUMSET
	1=Top set for maths	
	2=Middle set for maths	
	3=Bottom set for maths	
Number of classes in child's school year at MCS 4	Range 1 to 18	dq2524

Variable name	Description	MCS Variable(s) used
Number of classes in child’s school year at MCS 5	Range 1 to 10	EQ35
Number of children within child’s classroom at MCS 4	Range 1 to 48	dq2511
Number of children within child’s classroom at MCS 5	Range 1 to 64	EQ33
Number of children with English as additional language in child’s classroom at MCS 4	Range 0 to 39	dq2524
Number of children with English as additional language in child’s classroom at MCS 5	Range 0 to 88	EQ38

Variable name	Description	MCS Variable(s) used
Child's class contain a mixed year group at MCS 4	0=No	dq2513
	1=Yes	
Child's class contain a mixed year group at MCS 5	0=No	EQ34
	1=Yes	

Table A3. 6: Sensitivity analysis: The association between transition to attainment grouping and changes in children's parent- and teacher-reported externalising psychological symptoms between MCS 4 and MCS 5 by verbal skills at MCS 3 (adjusted for transition to top, middle and bottom streams/sets)

	Parent-reported externalising symptoms				**Teacher-reported externalising symptoms**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change from MCS 4 to MCS 5	-0.24**	0.08	-0.40	-0.08	0.10	0.12	-0.14	0.33
Attainment grouping * MCS5	-0.11	0.19	-0.47	0.26	-0.15	0.40	-0.93	0.63
Transition to streams (Ref. Mixed)								
Top	-0.13	0.12	-0.36	0.11	-0.04	0.19	-0.41	0.32
Middle	0.02	0.16	-0.30	0.33	-0.12	0.29	-0.68	0.44
Bottom	-0.11	0.24	-0.58	0.36	0.40	0.53	-0.64	1.43
Transition to literacy set groups (Mixed)								
Top	-0.07	0.10	-0.27	0.13	0.33	0.17	0.00	0.66
Middle	-0.14	0.13	-0.39	0.11	0.09	0.23	-0.37	0.55
Bottom	0.14	0.17	-0.20	0.48	0.23	0.35	-0.45	0.90

	Parent-reported externalising symptoms				Teacher-reported externalising symptoms			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Transition to maths set groups (Ref. Mixed)								
Top	0.08	0.17	-0.25	0.41	-0.17	0.38	-0.91	0.57
Middle	0.27	0.18	-0.09	0.63	-0.12	0.40	-0.92	0.67
Bottom	0.17	0.20	-0.23	0.58	0.20	0.46	-0.70	1.10

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** *p<0.001, ** <0.01, * <0.05*

Table A3. 7: Sensitivity analysis: The association between transition to attainment grouping and changes in children's parent- and teacher-reported internalising psychological symptoms between MCS 4 and MCS 5 by verbal skills at MCS 3 (adjusted for transition to top, middle and bottom streams/sets)

	Parent-reported internalising symptoms				**Teacher-reported internalising symptoms**			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Change form MCS 4 and MCS 5	0.54***	0.08	0.38	0.71	0.09	0.102	-0.11	0.29
Attainment grouping * MCS5	-0.18	0.21	-0.59	0.23	0.03	0.258	-0.48	0.54
Transition to streams (Ref. Mixed)								
Top	-0.15	0.13	-0.40	0.09	0.01	0.154	-0.30	0.31
Middle	-0.05	0.16	-0.37	0.26	-0.02	0.206	-0.42	0.39
Bottom	-0.02	0.26	-0.53	0.49	0.15	0.319	-0.47	0.78
Transition to literacy set groups (Ref. Mixed)								
Top	-0.05	0.11	-0.26	0.17	0.05	0.137	-0.22	0.32
Middle	-0.19	0.13	-0.44	0.06	-0.01	0.177	-0.36	0.34

	Parent-reported internalising symptoms				Teacher-reported internalising symptoms			
VARIABLES	**B**	**(SE)**	**95% CI**	**95% CI**	**B**	**(SE)**	**95% CI**	**95% CI**
Bottom	-0.10	0.18	-0.46	0.25	0.20	0.244	-0.28	0.68
Transition to maths set groups (Ref. Mixed)								
Top	-0.04	0.18	-0.40	0.32	-0.25	0.232	-0.71	0.21
Middle	0.09	0.19	-0.29	0.47	-0.24	0.248	-0.73	0.25
Bottom	0.64**	0.22	0.20	1.09	0.15	0.277	-0.39	0.70

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Appendix 4

Table A.4 1:Variable labels and definitions for empirical analysis in chapter 7.

Variable name	**Description**	**MCS Variable(s) used**
Key Explanatory variable		
Attainment grouping at MCS 4	The 3 variables were used to derive the attainment grouping status where 0=Not grouped by attainment, 1=Grouped by attainment (either for streaming or literacy setting or maths setting)	streamg, litsetg, numsetg
Outcome(s)		
Parent-reported externalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	ddhypea0+ddconda0
Parent-reported externalising SDQ at MCS 5	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	EDHYPEA00+EDCOND00

Variable name	Description	MCS Variable(s) used
Parent-reported externalising SDQ at MCS 6	A variable was derived from derived SDQ subscales concerning Hyperactivity and Conduct problem behaviour symptoms. A higher score indicates higher externalising problem behaviour symptoms (Scores range from 0 to 20).	FCONDUCT+FHYPER
Parent-reported internalising SDQ at MCS 4	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	dpeer00+demot00
Parent-reported internalising SDQ at MCS 5	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	EDEMOT00+EDPEER00
Parent-reported internalising SDQ at MCS 6	A variable was derived from derived SDQ subscales concerning peer relationship and emotional problem behaviour symptoms. A higher score indicates higher internalising problem behaviour symptoms (Scores range from 0 to 20).	FEMOTION+FPEER

Variable name	Description	MCS Variable(s) used
School-related characteristics at MCS 4		
School type (parent-reported)	The 2 variables were used to derive a school type variable where 0=state maintained primary school, 1=fee-paying primary school and 2=faith primary school	dmsctya0+dmfthsa0
Number of classes in CMs year at MCS 4 (teacher-reported)	Number of classes range from 1 to 18	dq2524
Number of pupils with SEN in CM's class at MCS 4(teacher-reported)	Number of pupils with SEN in CM's class range from 0 to 21	dq2526
Number of pupils with EAL in CM's class at MCS 4(teacher-reported)	Number of pupils with SEN in CM's class range from 0 to 39	dq2530
Disruptive pupils in CM's class at MCS 4 (teacher-reported)	0=no disruptive pupils, 1=disruptive pupils in CMs class	dq2532
Class size (teacher-reported)	Number of pupils in CM's class range from 1 to 48	dq2511
Years of teaching experience at MCS 4 (teacher-reported)	Years that teachers have taught altogether range from 1 to 48	dq2481

Variable name	Description	MCS Variable(s) used
CMs class contains mixed year group at MCS 4 (teacher-reported)	0=no, 1=yes	dq2513
Survey design variable		
MCS strata	0=England-Advantaged, 1=England-Disadvantaged, 2=England-Ethnic, 3=Wales-Advantaged, 4=Wales-Disadvantaged, 5=Scotland-Advantaged, 6=Scotland-Disadvantaged, 7=Northern Ireland-Advantaged, 8=Northern Ireland-Disadvantaged	PTTYPE2
Effect modifiers		
BAS naming vocabulary (age 5)	Scores range from 20 to 80. A higher score indicates a higher verbal cognitive ability	cdnvtscr
Family income	Equivalised family income quintiles where 0=Highest incomes, 1=2nd quantile, 2=3rd quantile, 3=4th quantile, 4=Lowest incomes	DOECDUK00

Table A.4 2: Sensitivity analysis; Fixed and random effects for Model 2; the association between attainment grouping at age 7, and parent-reported externalising and internalising symptoms from 7 to 14 years (excluding the Scottish and Northern Irish MCS sample)

	Externalising symptoms				Internalising symptoms			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.14	0.13	-0.11	0.39	0.19	0.1	-0.01	0.38
Age	-0.14***	0.04	-0.21	-0.06	0.57***	0.04	0.49	0.65
Attainment grouping at age 7* age	-0.04	0.06	-0.16	0.07	-0.05	0.06	-0.17	0.07
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.70***	0.13	0.44	0.96	0.51***	0.11	0.3	0.72
England - Ethnic	0.48	0.24	0.00	0.95	0.47*	0.2	0.08	0.86
Wales - Advantaged	-0.07	0.27	-0.60	0.46	0.06	0.22	-0.37	0.49
Wales - Disadvantaged	0.06	0.22	-0.38	0.49	0.24	0.18	-0.12	0.59
Female	-1.24	0.11	-1.46	-1.01	-0.05	0.09	-0.23	0.13
School type (Ref. State-maintained)								
Fee-paying	-1.29***	0.28	-1.84	-0.75	-1.08***	0.23	-1.52	-0.63
Faith	-0.50***	0.13	-0.76	-0.24	-0.29*	0.11	-0.5	-0.08
Classroom size	-0.04***	0.01	-0.06	-0.02	-0.03*	0.01	-0.05	-0.01
Disruptive peers in CM's class	0.38**	0.12	0.14	0.63	0.36***	0.1	0.16	0.56

	Externalising symptoms				Internalising symptoms			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Number of pupils with SEN in CM's class	0.06*	0.02	0.01	0.11	0.01	0.02	-0.03	0.05
Number of pupils with EAL in CM's class	0.00	0.01	-0.02	0.02	0.01	0.01	-0.01	0.02
Random-effects Parameters								
Between-child intercept variance	3.35***	0.08	3.21	3.51	1.98***	0.1	1.8	2.18
Between-child slope variance	0.94***	0.04	0.87	1.02	0.84***	0.05	0.75	0.94
Between-child intercept and slope covariance	-0.53***	0.03	-0.58	-0.47	-0.23***	0.07	-0.37	-0.09
Between-occasion variance	1.73***	0.02	1.68	1.78	1.85***	0.02	1.81	1.9

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

*** p<0.001, ** <0.01, * <0.05*

Table A.4 3: Sensitivity analysis; Fixed and random effects for Model 3; the association between attainment grouping at age 7, family income and parent-reported externalising and internalising symptoms from 7 to 14 years (excluding the Scottish and Northern Irish MCS sample)

	Externalising symptoms				**Internalising symptoms**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.03	0.19	-0.35	0.42	0.27	0.16	-0.05	0.58
Age	-0.10	0.07	-0.25	0.04	0.52***	0.07	0.37	0.66
Attainment grouping at age 7 * age	-0.15	0.12	-0.38	0.08	-0.09	0.11	-0.32	0.13
Family income (Ref. Highest incomes								
2nd quantile	0.07	0.15	-0.23	0.36	0.28*	0.14	0.02	0.55
3rd quantile	0.16	0.16	-0.15	0.48	0.52***	0.14	0.25	0.80
4th quantile	0.81***	0.18	0.46	1.17	0.87***	0.16	0.56	1.17
Lowest incomes	0.89***	0.21	0.47	1.30	1.01***	0.17	0.67	1.35
Family income (Ref. Highest incomes* age								
2nd quantile *age	0.03	0.24	-0.44	0.49	-0.03	0.11	-0.24	0.19
3rd quantile * age	0.33	0.25	-0.16	0.82	0.01	0.11	-0.21	0.23

	Externalising symptoms				Internalising symptoms			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
4th quantile *age	-0.25	0.28	-0.79	0.30	0.27*	0.13	0.03	0.52
Lowest incomes * age	0.38	0.31	-0.24	0.99	0.27*	0.14	0.00	0.54
Attainment grouping at age 7* Family income (Ref. Highest incomes)								
Attainment grouping at age 7 * 2nd quantile	-0.02	0.11	-0.25	0.20	0.08	0.21	-0.34	0.50
Attainment grouping at age 7 * 3rd quantile	0.03	0.12	-0.19	0.26	-0.51*	0.22	-0.95	-0.08
Attainment grouping at age 7 * 4th quantile	-0.05	0.13	-0.31	0.21	-0.10	0.24	-0.56	0.37
Attainment grouping at age 7 * Lowest incomes	-0.02	0.15	-0.31	0.27	0.17	0.26	-0.33	0.67
Attainment grouping at age 7* Family income (Ref. Highest incomes) * age								
Attainment grouping at age 7 * 2nd quantile *age	0.19	0.18	-0.16	0.54	-0.04	0.17	-0.38	0.30
Attainment grouping at age 7 * 3rd quantile*age	-0.01	0.18	-0.37	0.35	0.52**	0.18	0.17	0.87

	Externalising symptoms				Internalising symptoms			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Attainment grouping at age 7 * 4th quantile * age	0.48	0.20	0.08	0.88	0.01	0.19	-0.37	0.39
Attainment grouping at age 7 * Lowest incomes *age	-0.08	0.22	-0.52	0.35	-0.40	0.21	-0.80	0.01
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.46***	0.13	0.20	0.72	0.28*	0.10	0.08	0.48
England - Ethnic	0.16	0.24	-0.31	0.63	0.22	0.18	-0.14	0.57
Wales - Advantaged	-0.09	0.26	-0.61	0.42	-0.05	0.21	-0.45	0.35
Wales - Disadvantaged	-0.13	0.22	-0.56	0.29	-0.04	0.16	-0.36	0.27
Scotland - Advantaged								
Scotland - Disadvantaged								
Northern Ireland - Advantaged								
Northern Ireland - Disadvantaged								
Female	-1.25***	0.11	-1.47	-1.03	-0.1	0.083625	-0.26674	0.061066
School type (Ref. State-maintained)								
Fee-paying	-0.99***	0.27	-1.39	-0.45	-0.66**	0.21	-1.07	-0.24
Faith	-0.46***	0.13	-0.71	-0.21	-0.27**	0.10	-0.46	-0.07
Classroom size	-0.03**	0.01	-0.05	-0.01	-0.02*	0.01	-0.03	0.00
Disruptive peers in CM's class	0.36**	0.12	0.12	0.60	0.27**	0.09	0.09	0.44
Number of pupils with SEN in child’s class	0.05	0.02	0.01	0.10	0.02	0.02	-0.01	0.05
Number of pupils with EAL in child’s class	0.00	0.01	-0.03	0.02	0.01	0.01	-0.01	0.02
Random-effects Parameters								

	Externalising symptoms				Internalising symptoms			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Between-child intercept variance	3.27***	0.08	3.12	3.42	1.94***	0.10	1.76	2.14
Between-child slope variance	0.94***	0.04	0.86	1.03	0.80***	0.05	0.71	0.91
Between-child intercept and slope covariance	-0.54***	0.03	-0.59	-0.48	-0.21***	0.08	-0.36	-0.05
Between-occasion variance	1.78***	0.02	1.73	1.82	1.88***	0.02	1.83	1.93

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*

Table A.4 4: Sensitivity analysis; Fixed and random effects for Model 4; the association between attainment grouping at age 7, initial verbal skills and parent-reported externalising and internalising symptoms from 7 to 14 years (excluding the Scottish and Northern Irish MCS sample)

	Externalising symptoms				**Internalising symptoms**			
	B	**SE**	**95% CI**	**95% CI**	**B**	**SE**	**95% CI**	**95% CI**
Fixed-effects Parameters								
Attainment grouping at age 7	0.48	0.62	-0.73	1.69	0.88	0.49	-0.07	1.84
Age	-0.05	0.20	-0.45	0.35	1.04***	0.21	0.64	1.45
Attainment grouping at age 7 * age	-0.54	0.30	-1.13	0.06	-0.49	0.31	-1.09	0.11
Verbal skills at age 5	-0.06***	0.01	-0.07	-0.04	-0.02***	0.01	-0.03	-0.01
Verbal skills at age 5 * age	0.00	0.00	-0.01	0.01	-0.01*	0.00	-0.02	0.00
Attainment grouping at age 7*Verbal skills at age 5	-0.01	0.01	-0.03	0.02	-0.01	0.01	-0.03	0.00
Attainment grouping at age 7 * verbal skills at age 5 * age	0.01	0.01	0.00	0.02	0.01	0.01	0.00	0.02
MCS stratum (Ref. England-Advantaged)								
England - Disadvantaged	0.60***	0.12	0.36	0.85	0.47***	0.10	0.27	0.67
England - Ethnic	0.15	0.23	-0.29	0.59	0.45*	0.19	0.09	0.82
Wales - Advantaged	-0.27	0.26	-0.77	0.23	-0.05	0.21	-0.46	0.37
Wales - Disadvantaged	-0.16	0.20	-0.55	0.22	0.17	0.16	-0.15	0.48
Female	-1.18***	0.10	-1.38	-0.97	-0.04	0.09	-0.21	0.13

School type (Ref. State-maintained)								
Fee-paying	-0.87**	0.26	-1.38	-0.36	-0.81***	0.21	-1.23	-0.39
Faith	-0.41**	0.12	-0.64	-0.17	-0.26*	0.10	-0.45	-0.06
Classroom size	-0.02	0.01	-0.04	0.00	-0.01	0.01	-0.03	0.00
Disruptive peers in CM's class	0.28*	0.11	0.06	0.51	0.26*	0.09	0.07	0.44
Number of pupils with SEN in child's class	0.04*	0.02	0.00	0.09	0.02	0.02	-0.02	0.05
Number of pupils with EAL in child's class	-0.01	0.01	-0.03	0.01	0.00	0.01	-0.02	0.01
Constant	9.80***	0.53	8.77	10.83	4.02***	0.42	3.19	4.86
Random-effects Parameters								
Between-child intercept variance	3.25***	0.08	3.10	3.40	2.01***	0.10	1.83	2.20
Between-child slope variance	0.93***	0.04	0.86	1.02	0.84***	0.05	0.75	0.94
Between-child intercept and slope covariance	-0.53***	0.03	-0.58	-0.47	-0.25***	0.07	-0.38	-0.11
Between-occasion variance	1.76***	0.02	1.72	1.81	1.86***	0.02	1.82	1.91

Ref. stands for refence category

Stars indicate that the difference is significant according to two-tailed *t*-tests.

**** p<0.001, ** <0.01, * <0.05*